Interferometric Synthetic Aperture Sonar Design and Performance

Philip J. Barclay, B.E. (Hons)

A thesis presented for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering
at the
University of Canterbury,
Christchurch, New Zealand.

August 2006

ABSTRACT

Synthetic aperture sonar (SAS) has become a well developed imaging technique for imaging shallow water environments. Aperture synthesis provides high along-track resolution imagery, with range independent resolution. However, mapping of the seafloor using traditional SAS is limited to a two-dimensional surface. To provide the third dimension (height), an interferometric synthetic aperture sonar (InSAS) is formed, comprising of two or more vertically displaced hydrophone arrays. Each of the interferometric receiver datasets are processed using standard SAS algorithms, with motion compensation and corrective processing applied equally to each channel, preserving the underlying interferometric time delays. By then estimating the time delay of the incoming wavefronts across the interferometric receiver array, the height of the seafloor can be inferred from the side-scan geometry of the system. The InSAS approach is similar to the radar equivalent (InSAR), however, significant differences in geometry and medium properties limit the applicability of InSAR algorithms to the sonar equivalent.

A height estimate from interferometric data is formed by estimating the time difference between the receiver elements of the interferometric array. Therefore, for an accurate estimate of the time-delay, the signals of the receivers must contain significant 'common' information. Presented in this thesis is an analysis of coherence as applicable to an InSAS system. The coherence of an InSAS system can be decomposed into five 'coherence components': additive acoustic noise, footprint misalignment, baseline decorrelation, temporal decorrelation, and processing noise. Of these, it is shown footprint misalignment has the greatest effect for an InSAS system if it is not corrected for. The importance of maintaining high coherence between the receiver channels is presented; small losses in coherence from the ideal of unity will have a significant impact of the accuracy of the resulting height estimate. To reduce the sensitivity of the height accuracy losses, multiple estimates of the height can be formed from independent 'looks' of the scene. Combining all these estimates into one height estimate is shown to significantly improve the height estimate.

The design and signal processing of an InSAS system is of high importance to the generation of high accuracy height estimates of the seafloor. Several parameters of design are explored, in particular the effect of aperture sampling. Low along-track aperture sampling rates are shown to cause a significant decrease in signal coherence,

caused by the generating of 'grating lobes' from the synthetic aperture processing. Substantial improvements can be made by careful selection of transmitter and receiver element sizes, relaxing the requirements of a highly sampled aperture. An analysis of interpolation schemes on interferometric quality is also presented.

The effect of footprint misalignment can be reduced by first resampling the data from each receiver onto a common ground-plane. However, this requires prior knowledge of the seafloor height, an unknown parameter before an interferometric height estimate is made. One possible method to form an initial height estimate is through the use of belief propagation, a technique applied from the field of stereo imaging. Belief propagation is used to estimate an initial height surface, albeit at discrete height intervals. This initial low resolution height surface can then be used to remap the data, partially eliminating the detrimental effects of footprint misalignment.

The combination of all the independent estimates of the scene can be combined using maximum likelihood estimation. This framework allows the individual estimates to be combined into one overall cost function. Searching of the cost function for minimum cost yields a single interferometric time-delay estimate, from which a single height estimate can be inferred. This framework allows looks formed from many different sources to be combined, including multiple imaging frequency bands, and the use of more than one interferometric pair of receivers.

ACKNOWLEDGEMENTS

Thank you to everyone who has helped me throughout this thesis. To my supervisors, Dr. Michael Hayes and Prof. Peter Gough, thank you for your guidance and useful discussions along the way. Thank you to the rest of the Acoustic Research Group, in particular Hayden Callow, Steve Fortune, Alan Hunter, and Edward Pilbrow. Also, thanks to Chris Forne for help with the application of belief propagation algorithms.

During my stay as a postgrad in the ECE department I have had much support from the department, both from students and staff. It is these fantastic people that make writing a thesis possible. The construction and operation of the KiwiSAS sonar system was an integral part of the early work conducted in this thesis. Much of the mechanical work for the KiwiSAS system was performed by the technical staff of the ECE Department, in particular thanks goes to Michael Cusdin, Peter Lambert, and Steve Downing. For sea-trials of the KiwiSAS system, the use of Peter's yacht was of much help, I hope we didn't scratch it too much dragging things back onboard.

To my family, thank you for always believing in me, and being there for me through the hard times. A special thank you to Helen. Your love and support means a lot to me.

Thank you also to all my friends, especially those who dragged me away into the hills to get away from it all for a while.

PREFACE

The work in this thesis originally began as a hardware based extension of the KiwiSAS-II sonar system to allow for interferometric operation. The idea was to re-wire the existing 3×3 array of hydrophone elements into three vertically separated receivers, allowing interferometry to be performed. For this, the acquisition hardware was extended to six channels, allowing each hydrophone to operate within two separate frequency bands, centered on 30 kHz and 100 kHz. The hardware developed for this system is covered in [Barclay et al., 2002a], expanded in Appendix B. Sea-trials with this hardware were performed locally in Lyttelton Harbour [Barclay et al., 2001], and the Hauraki Gulf Auckland [Hayes et al., 2001].

Later, the hardware was completely redesigned and rebuilt since it was apparent additional hydrophones would be of benefit. Instead of sending the received waveforms up the cable in an analogue form, all the conversion and storage electronics was moved into the towfish housing. This was necessitated by the large increase in overall datarate from the expanded sonar configuration. Central to this electronics package is a Pentium-III industrial computer, mounted within a Compact-PCI cage [Hayes et al., 2002]. Waveform generation and echo recording implemented as three custom Compact-PCI cards. This configuration allows the transmitter to be driven in vertical groups of three, allowing synthetic horizontal beam shaping. The new receiver hardware provides nine channels, allowing the 3×3 receiver array to be individually sampled, each with a bandwidth of $\approx$150 kHz. Post-processing is used to extract the two 20 kHz bandwidth signals.

During initial interferometric processing, direct phase differences between the receiver channels were used as the basis of the angle-of-arrival [Barclay et al., 2002b, Barclay et al., 2003b]. Direct phase differences between the receivers was found to give poor results; the coherence loss due to footprint shift being the primary problem. This is covered in more detail in Chapter 3. The coherence loss due to footprint shift was exploited by resampling the data onto a range of differing height ground-planes, with a surface mapped through the solution space. The fitting of the surface through the solution space was performed using belief propagation, a technique first developed in the computer graphics field. This technique provides a good first order approximation to the seafloor topography. The details of this application of belief propagation can be

found in [Barclay et al., 2003a], reproduced in Appendix A.

After a more in-depth analysis of signal coherence between the receivers was performed (see Chapter 3), it became apparent multiple 'looks' of the scene would be required to provide a high accuracy height estimate. Through various techniques detailed in Chapter 4, multiple looks of the scene could be generated. However, the combination of all the looks into one overall height estimate proved difficult, especially when combining data from the two distinct frequency bands (30 and 100 kHz). To achieve this combination, a maximum likelihood approach was taken. The data from each look is formed such that a single parameter search can be made to find the most likely angle-of-arrival. This approach has shown to be highly successful, results presented in several papers [Barclay et al., 2004, Barclay et al., 2005, Barclay et al., 2006]. One major problem with a maximum likelihood approach is the assumption of a single angle-of-arrival, an assumption violated by multipath from the sea-surface. Work on this problem has also been performed both using belief propagation [Hayes and Barclay, 2003] and maximum likelihood approaches [Barclay et al., 2006].

Papers prepared during this thesis are presented here in order of presentation.

Barclay, P. J., Forne, C. J., Hayes, M. P., and Gough, P. T. (2003a). Reconstructing seafloor bathymetry with a multi-frequency, multi-channel broadband InSAS using belief propagation. In *Oceans 2003*. IEEE, MTS.

Barclay, P. J., Hayes, M. P., and Gough, P. T. (2001). Using a multi-frequency synthetic aperture sonar for bathymetry. In *Image and Vision Computing New Zealand 2001*, pages 63–68, University of Otago, Dunedin, New Zealand.

Barclay, P. J., Hayes, M. P., and Gough, P. T. (2002a). Multi-channel data acquisition for a free-towed synthetic aperture sonar. In *Proceedings of Electronics New Zealand Conference*, pages 133–138, Dunedin, New Zealand.

Barclay, P. J., Hayes, M. P., and Gough, P. T. (2002b). Reconstructing seafloor bathymetry with a multichannel broadband InSAS. In *Image and Vision Computing New Zealand 2002*, pages 47–52, University of Auckland, Auckland, New Zealand.

Barclay, P. J., Hayes, M. P., and Gough, P. T. (2003b). Bathymetry reconstruction for a free-towed synthetic aperture sonar. In *Proceedings of the World Congress on Ultrasonics*, pages 519–522.

Barclay, P. J., Hayes, M. P., and Gough, P. T. (2004). Multi-receiver, multi-frequency maximum likelihood synthetic aperture sonar interferometry. In *Proceedings of Image and Vision Computing New Zealand 2004*, pages 273–278, Akaroa, New Zealand.

Barclay, P. J., Hayes, M. P., and Gough, P. T. (2005). ML estimation of seafloor topography using multi-frequency synthetic aperture sonar. In *Oceans 2005 Europe*, Brest, France. IEEE.

Barclay, P. J., Hayes, M. P., and Gough, P. T. (2006). Bathymetric results from a multi-frequency InSAS sea-trial. In *Oceans 2006 Asia-Pacific*, Singapore. IEEE, OES.

Hayes, M. P. and Barclay, P. J. (2003). The effects of multipath on a bathymetric synthetic aperture sonar using belief propagation. In *Proceedings of Image and Vision Computing New Zealand, IVCNZ2003*, Palmerston North, New Zealand.

Hayes, M. P., Barclay, P. J., Gough, P. T., and Callow, H. J. (2001). Test results from a multi-frequency bathymetric synthetic aperture sonar. In *Oceans 2001*, volume 1, pages 1682–1688. IEEE.

Hayes, M. P., Barclay, P. J., and Hawkins, T. J. (2002). An embedded compact PCI computer system for a synthetic aperture sonar towfish. In *Proceedings of Electronics New Zealand Conference*, pages 139–144, Dunedin, New Zealand.

Hayes, M. P., Hunter, A. J., Barclay, P. J., and Gough, P. T. (2005). Estimating layover in broadband synthetic aperture sonar bathymetry. In *Oceans 2005 Europe*, Brest, France. IEEE.

DEFINITIONS

Imaging system

towboat tow vehicle for a towed system (not present for an autonomous system)

towfish sonar housing containing at least the transmitter and receiver elements

ping sound energy signal projected into water column

ping rate repetition rate of projected pings into water column

pulse (ping) repetition rate see ping rate

pulse (ping) repetition frequency frequency of projected pings, reciprocal of ping rate

resolution cell shape and size of resolvable patch of sea-floor

mapping rate rate of seafloor mapping coverage, normally expressed in square metres per hour

height-map estimated height of scene

Imaging geometry

across-track direction perpendicular to nominal direction of travel of sonar

along-track direction parallel to nominal direction of travel of sonar

multipath imaging paths with reflections from the sea-surface and/or sea-floor of total length equal to the direct path

bathymetry estimation of seafloor depth at a given point

footprint shift/mis-alignment across-track offset of resolution cells of different transmitter/receiver pairs

nadir direction directly down from sonar

zenith direction directly up from sonar

azimuth angular distance on horizontal-plane measured from directly in front of sonar to directly behind sonar

elevation angular distance on vertical plane from zenith and nadir

bore-sight direction directly away horizontally in across-track from sonar

x,y,z 3-dimensional coordinate frame, x in across-track direction, y in along-track direction, z upwards from sea-surface

pitch angular sonar motion around the x-axis (nose up/down)

roll angular sonar motion around the y-axis

yaw angular sonar motion around the z-axis (nose left/right)

side-scan geometry where imaging is performed primarily from the side of the platform

sway translational sonar motion along the x-axis (across-track direction)

surge translational sonar motion along the y-axis (along-track direction)

heave translational sonar motion along the z-axis (up/down direction)

Signal parameters

centre frequency central frequency of imaging signal

bandwidth spectral width of imaging signal centered on center frequency

sampling frequency frequency of signal sampling

Nyquist sampling frequency minimum sampling frequency required to adequately sample a band-limited signal without aliasing

quality factor, Q ratio of center frequency to bandwidth of signal

narrowband low bandwidth to center frequency ratio, typically $Q > 2$

broadband high bandwidth to center frequency ratio, typically $Q < 2$

Data processing

spatial changing over distance

temporal changing over time

base-banding spectral signal shifting from modulated center frequency to zero frequency

pulse-compression reduction of signal time extent by correlation with expected signal

motion estimation estimation of the motion of the sonar, either from the data itself, or from some external means (e.g., navigation unit)

auto-focus image focusing/sharpening of sonar imagery using only the data as information source

grating-lobe along-track artifact of synthetic aperture processing

ACRONYMS

AASR	Along-track ambiguity to signal ratio
ADC	Analogue to digital converter
AUV	Autonomous underwater vehicle
BNC	Coaxial connector (Bayonet Neill Concelman) or Bayonet Naval Connector
CAD/CAC	Computer aided detection / Computer aided classification
CARABAS	Coherent all radio band sensing
CAATI	Computed angle-of-arrival transient imaging
CPLD	Complex programmable logic device
CPU	Central processing unit
CRLB	Cramér-Rao lower bound
CSDE	Coherent source direction estimation
DAC	Digital to analogue converter
DDC	Digital down converter
DIN	Deutsches Institut für Normung eV (in English, the German Institute for Standardization)
DPIA	Displaced ping imaging auto-focus
DSP	Digital signal processor
DVL	Doppler velocity log
ERS-1,2	European remote sensing satellites 1&2
ESA	European space agency

FFBP	Fast factorised back projection
FFI	Norwegian Defence Research Establishment
FFT	Fast Fourier transform
FM	Frequency modulated
FOI	Swedish National Defence Research Establishment, Linköping, Sweden
FPBP	Fast polar back projection
FPGA	Field programmable gate array
GPS	Global positioning system
HF	High frequency
IDE	Integrated drive electronics
IEEE	Institute of Electrical and Electronic Engineers
InSAS	Interferometric synthetic aperture sonar
InSAR	Interferometric synthetic aperture radar
JPL	Jet propulsion laboratories
KiwiSAS	Kiwi synthetic aperture sonar (University of Canterbury)
LAM-MPI	Local area multicomputer - Message passing interface
laser	light amplification by stimulated emission of radiation
LFM	Linear frequency modulated
MATLAB	Matrix Laboratory (www.mathworks.com)
MBES	Multi-beam echo sounder
ML	Maximum likelihood
MRF	Markov random field
NMEA	National Marine Electronics Association
PCI	Peripheral component interconnect
PDF	Probability density function
PLL	Phase locked loop
PRF	Pulse repetition frequency
PRR	Pulse repetition rate

PROM	Programmable read-only memory
PSF	Point spread function
PVDF	Polyvinylidene fluoride
radar	radio detecting and ranging
RELAX	Relaxation algorithm
RMS	Root mean square
SAMI	Synthetic aperture mapping and imaging sonar
SA	Synthetic aperture
SAR	Synthetic aperture sonar
SAS	Synthetic aperture sonar
SCR	Signal to clutter ratio
SNR	Signal to noise ratio
sonar	sound navigation and ranging
SRC	Secondary range compression
TDM	Time division multiplexed
TRAR	Transmitter to receiver along-track size ratio
URI	Unambiguous range interval
VHF	Very high frequency

CONTENTS

Chapter 1

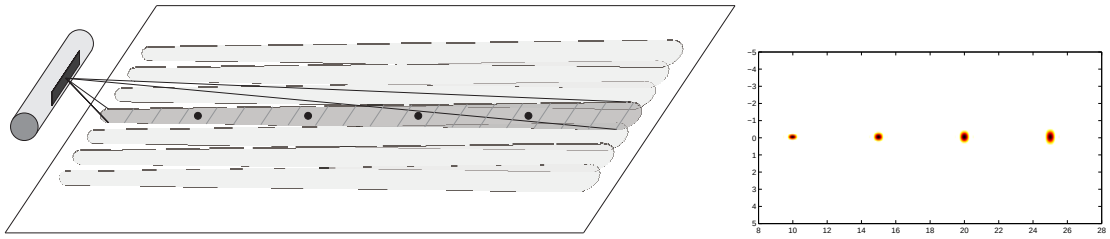
INTRODUCTION

The concept of synthetic aperture sonar (SAS) was primarily driven by the desire for high resolution imagery of the seafloor, with high mapping rates. In shallow water environments, high mapping rates require the use of long-range, side-looking sonar systems. The need for long range imaging requires the use of low frequencies, since the absorption of acoustic energy in water increases with frequency [Urick, 1975]. Using a conventional sonar system with real transmitter and hydrophone elements, the desire for high resolution at long ranges with low frequency signals implies the need for very long (impractically long) acoustic transducer arrays. To circumvent this problem, the concept of a 'synthetic' aperture is required; the large aperture required synthesised by moving a wide-beamwidth, small aperture along the imaged scene. Via coherent addition (within the SAS algorithms), these wide-beamwidth images produce an image with range independent resolution [Hawkins, 1996]. The differences between a real-aperture and synthetic-aperture system can be seen pictorially in Figure 1.1.

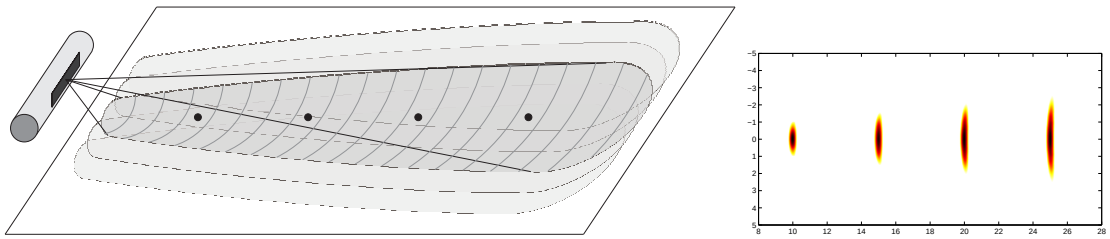
Interferometric synthetic aperture sonar (InSAS) is an extension to SAS; the goal to produce a height map to complement the conventional SAS backscattered intensity imagery. The basic principle of InSAS is to use a minimum of two receivers with vertical separation between them. The seafloor is illuminated with sound energy as would be the case for conventional SAS, the reflected echoes received by the interferometric pair of receivers. Each of the receivers is processed individually using standard phase-preserving SAS techniques, giving the starting point for InSAS algorithms. The time difference of received echoes across the interferometric pair is estimated, and via geometry, the height of the seafloor can be inferred. Since InSAS is an extension to SAS, a basic introduction to SAS operation and processing strategies is now presented.

1.1 SYNTHETIC APERTURE SONAR PROCESSING

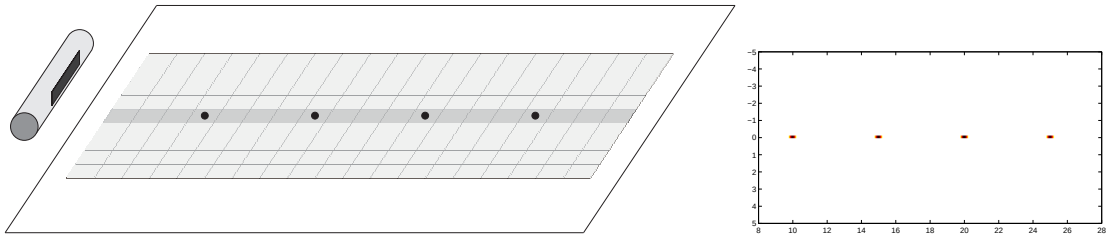
Every sonar system developed has different operational requirements, and hence different physical constructions. However, all of these systems have common sub-systems. The nature, size, count, physical location, and interconnection between these sub-



(a) Real aperture narrow-beam sonar.



(b) Synthetic aperture sonar before azimuth compression.



(c) Synthetic aperture sonar after azimuth compression.

Figure 1.1 Acoustic footprints of real and synthetic aperture sonars, and the resulting intensity imagery. A real aperture sonar (1.1(a)) utilises a narrow along-track beamwidth, giving poorer resolution at greater across-track ranges. A synthetic aperture sonar (1.1(b)) has a broad along-track beamwidth during data collection, giving a large change in along-track resolution with increasing across-track range. This is equivalent to a real aperture sonar with a large beam angle. After synthetic aperture processing (1.1(c)), the along-track resolution is improved, providing range independent resolution.

systems will differ for every system. The sub-systems common to all active SAS systems are: the transmitter(s) and associated electronics, the receiver(s) and associated electronics, a data storage sub-system, an overall control/data processing sub-system (often including human operators), and one or more electrical power supplies for each of these sub-systems. Although not essential for the most basic SAS systems, normally also present is a navigation sub-system comprised of many sensors. The transmitter, receiver, and navigation sub-systems are normally mounted on a common underwater body, while the location of the other sub-systems depends on the overall design of the system. Due to advances in digital electronics, the storage and control sub-systems are normally also located within the underwater body (often referred to as the 'towfish' in towed systems). The location of the power source varies depending on the platform type, normally located on the tow-vehicle for a towed system, and as part of the underwater body for an non-towed vehicle (such as an autonomous underwater vehicle (AUV)).

The data processing of the raw echoes received by the hydrophones is different for every sonar system according to the nature of the system, and the desired final imagery/data required. A typical processing chain is shown in Figure 1.2, but many variations will exist depending on the algorithms used. For a two dimensional data set as obtained from a SAS system, four Fourier processing domains exist as shown in Figure 1.3 for a single point target. The four domains are the: spatial-temporal domain, the temporal-frequency domain, the range-Doppler domain, and the wavenumber domain. Almost all algorithms associated with SAS processing utilise more than one of these domains, applying the Fourier transform to move between the various domains.

1.1.1 Data Collection

There are two main categories of SAS systems: towed platforms, and autonomous underwater vehicles (AUVs). A third category also exists, rail-based systems, but because of the non-portable nature of such systems, the capability of these systems is limited to research. AUV based SAS systems introduce many complications to the operation of the SAS sub-system, primarily due to the autonomous requirement of the platform. The autonomous nature of the vehicle implies the SAS must also work in an autonomous manner, with all data recorded within the vehicle itself for later downloading and analysis. AUVs must also be capable of both macro- and micro-navigation, requiring a multitude of sensors to be incorporated into the vehicle, often using the SAS imagery and motion estimation as a source of navigational information. By comparison, a towed system with a surface tow vehicle can utilise sensors on the tow vehicle for macro-navigation, simplifying the sonar system design. Towed platforms also often allow partial or complete access to the sonar imagery in real-time, allowing human operators on the tow vehicle to partially analyse the data in real-time allowing

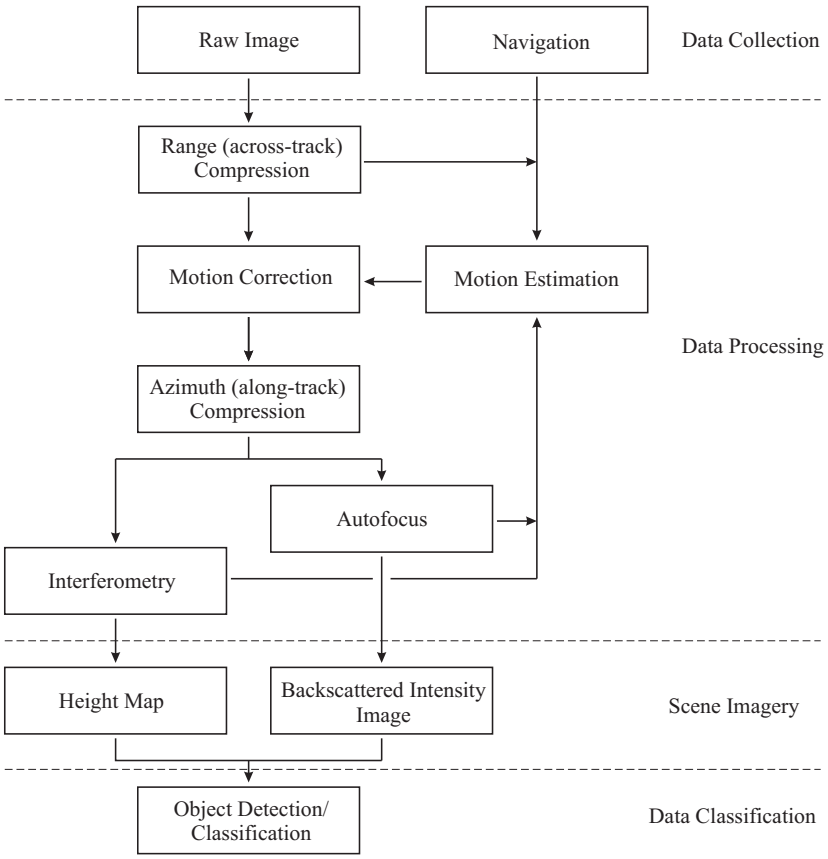
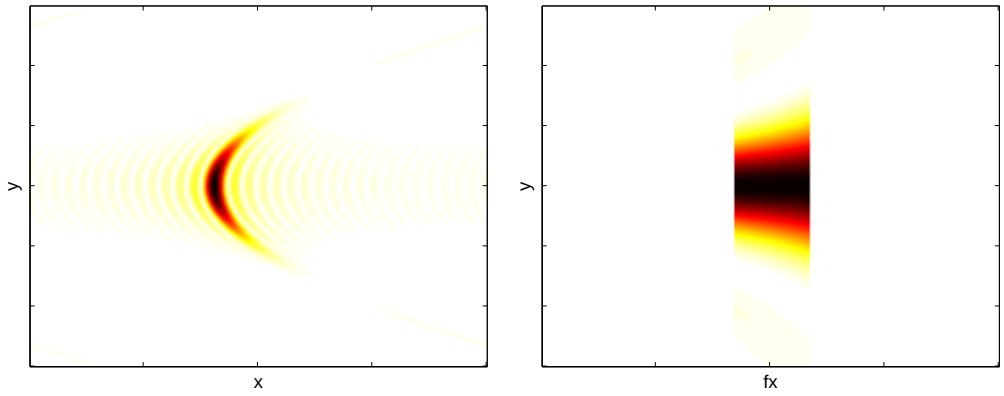
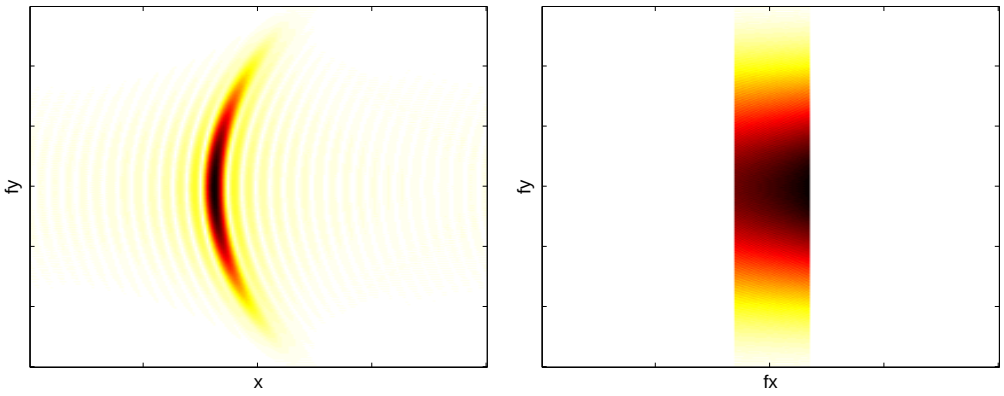


Figure 1.2 Flow diagram of the typical data processing sequence of a synthetic aperture sonar system. The data flow within the processing chain will deviate from this layout for certain algorithms. In particular, the motion estimation and correction steps can vary according to algorithm, often occurring in multiple locations of the processing chain.



(a) Spatial, temporal domain.

(b) Spatial, temporal frequency domain.



(c) Range Doppler domain.

(d) Wavenumber domain.

Figure 1.3 A single pulse-compressed point target in the four Fourier processing domains: Figure 1.3(a) Spatial-temporal domain, Figure 1.3(b) Spatial, temporal frequency domain, Figure 1.3(c) Range-Doppler domain, Figure 1.3(d) Wavenumber domain. Transformations between the four domains are performed with a one-dimensional Fourier transform in the appropriate direction.

alterations of the operational parameters of the sonar during data collection. While this initially seems an advantage, there are many disadvantages of using a towed system, especially when imaging large areas or potentially dangerous terrain. Large seafloor areas and practical time constraints imply the use of several SAS systems, requiring several tow vehicles and operators. If AUVs were used for a similar task, it would be possible deploy several vehicles from one surface vehicle from some safe distance from the potentially dangerous area of imaging, minimising the chances of the loss of human life. For research and development of SAS systems and algorithms, towed arrays minimise the cost and complexity of the sonar platform, while still providing real-world problems such as platform motion to be investigated.

1.1.2 Mapping and data rates

To increase the mapping rate while maintaining imaging resolution, along-track receiver arrays of hydrophones are often used. Using large receiver arrays with many tens or even hundreds of receiver elements produces a large quantity of data, increasing the storage and computing requirements for the subsequent image processing stages. Only recently have such large receiver arrays become feasible, due to advances in digital electronics and data storage. For example, consider a hypothetical sonar system with an along-track receiver array of $N_r = 64$ elements, a bandwidth of $B = 20\,\text{kHz}$ centered at $f_c = 100\,\text{kHz}$, with 16 bit ($N_w = 2$ bytes) sampling. Thus, the highest frequency signal present is 110 kHz, giving a Nyquist sampling rate of 220 kHz. Practically, a sampling rate of higher than this minimum is required to allow for spectral leakage and shaping, so a sampling rate of $f_s = 250\,\text{kHz}$ is chosen. Assuming the ping rate is chosen to give continuous sampling between pings, the raw data rate of this sonar is,

$$\begin{aligned} \text{datarate}_{\text{raw}} &= f_s \times N_w \times N_r \\ &= 250 \times 10^3 \times 2 \times 64 \\ &= 32\,\text{MB/s}. \end{aligned} \tag{1.1}$$

For one hour of operation this system will produce approximately 115 GB of raw data, requiring large and expensive storage sub-systems. Of the theoretical 125 kHz bandwidth, only a small fraction (20 kHz) contains active imaging information. By complex base-banding the data in real-time before storage and subsequent processing, the data volume can be greatly reduced. Thus, if the signal is basebanded to a sampling rate of $f_s = 20\,\text{kHz}$ with complex samples (requiring $N_w = 4$ bytes per sample), the data rate is reduced to,

$$\begin{aligned} \text{datarate}_{\text{baseband}} &= 20 \times 10^3 \times 4 \times 64 \\ &= 5.12\,\text{MB/s}. \end{aligned} \tag{1.2}$$

For an hour's worth of data collection, this now produces a more manageable 18.5 GB of data. However, it should be noted the computational processing power required to perform the down-conversion is high, with an input data rate of the original 32 MB/s. Extra care must be taken to ensure the accuracy of the base-banding algorithm if only the basebanded data is permanently stored. Any errors or data distortions introduced by the base-banding process will be difficult to reverse with post-processing.

Navigation sensors

Navigation sensors are also normally present on the data collection platform. The output of these sensors is processed for dual purposes, macro- and micro-navigation. Macro-navigation refers to the overall position of the towfish and imaged scene, allowing positions within the imagery to be defined on an absolute world scale. Macro-navigation is important for large area surveys where multiple sonar swaths are required to fully cover the scene, or surveys where objects of interest must be located accurately for later reference. For many research SAS systems, the accuracy of the macro-navigation is of less importance and often disregarded. Micro-navigation refers to the position and orientation of the sonar platform on a much smaller scale, but often has a requirement of higher accuracy than macro-navigation. Typically, the output of the micro-navigation system is used for motion correction of the imagery, ideally providing sub-wavelength accuracies over the length of the synthetic aperture. For more discussion on navigation sensors, see Section 5.5. Synchronization of the navigation sensors and the acoustic data is important, ensuring the motion corrections are applied to the appropriate portion of data. Many navigation sensors also have a lag between the motion occurring and being sensed, before motion estimation and correction is performed.

1.1.3 Across-track (range) compression

The across-track (range) resolution of a SAS system is determined by the bandwidth of the acoustic signal used. Most SAS systems use a time-varying frequency signal, commonly called a frequency modulated (FM) signal [Hawkins, 1996]. The use of a time-varying frequency signal allows the full bandwidth signal to be projected into the water column with greater overall energy. After echo collection from the hydrophones, the signal is matched filtered with the transmitted signal to reduce the time extent. This process is also known as pulse-compression. The signal spectrum may be further modified to improve across-track side-lobe suppression, covered in more detail in Section 3.5.4. For numerical efficiency, across-track compression is often performed in the spatial-frequency domain.

At the extreme across-track far range, the time-varying transmitted signal extends beyond the scene edge. When this portion of the scene is pulse-compressed with the transmitted reference signal, the result has a larger extent due to the reduction in

effective signal bandwidth. Thus, caution must be used when further processing this portion of the scene, and is often discarded before further processing.

1.1.4 Along-track compression: Synthetic aperture reconstruction algorithms

Synthetic aperture algorithms (SA) can be divided into four categories, dependent on the Fourier processing domain as shown in Figure 1.3. The first of these categories is the spatial-temporal domain, where algorithms process the data ordered as an image distance in both across- and along-track directions. The second domain is obtained by taking a Fourier transform in the across-track direction, giving the temporal-frequency domain. If the Fourier transform is instead performed in the along-track direction, the range-Doppler domain is formed. Finally, by Fourier transforming the spatial-temporal domain data in both along and across-track directions, the wavenumber domain is formed. Currently, there are five main synthetic aperture processing techniques: back-projection and correlation in the spatial-temporal domain, range-Doppler in the range-Doppler domain, and wavenumber and chirp-scaling in the wavenumber domain. Variations of many of the algorithms listed here exist, often utilising different image domains within the algorithm to gain numerical efficiency. For example, fast correlation methods exist utilising the spatial-frequency domain to provide numerical speedups for the correlations. For a comparison of these synthetic aperture algorithms see [Hawkins, 1996] and [Gough and Hawkins, 1997]. A brief summary of the main algorithms is presented here, with references for more detailed analysis.

Correlation method [Bamler, 1992, Chang et al., 1992] The correlation method processes the data in the spatial-temporal domain, used for many of the early digital synthetic aperture systems [Barber, 1985]. The direct correlation method is applied after across-track pulse-compression has been performed. The algorithm operates by correlating the expected point spread function (PSF) with the pulse-compressed data along the locus for a given point. Thus, for every point in the image, the pulse-compressed data must be interpolated to points along the locus of that point, multiplied by the complex conjugate of the PSF for that point, and summed to give the final result for that point. Since every point in the image has a differing locus path, the interpolation of the data is computationally expensive, making this algorithm extremely slow. A less computationally expensive version of this algorithm, known as fast-correlation, can be implemented using Fourier techniques [Hayes and Gough, 1992, Hawkins and Gough, 1995].

Back projection Back projection can be thought of as the geometric reverse of the forward imaging process, processing the data in the spatial-temporal domain. After

pulse-compression, each point in the image is 'back-projected' to all points in the reconstructed image of equal range. More recently back-projection has been developed into 'fast' equivalents, namely fast factorized back projection (FFBP) [Houston, 1994, Ulander et al., 2003, Banks and Griffiths, 2002, Banks, 2002] and fast polar back projection (FPBP) [Shippey et al., 2005]. These algorithms have been shown to have similar computational requirements as that of the traditionally faster wavenumber-domain algorithms such as wavenumber and chirp-scaling [Shippey et al., 2005].

Range-Doppler The range-Doppler algorithm was originally developed for space-borne SAR systems at the Jet Propulsion Laboratories (JPL) [Wu et al., 1982]. The algorithm operates in the range-Doppler domain, after the data has been pulse-compressed. Improvements were later made by [Jin and Wu, 1984], with the introduction of secondary range compression (SRC). For further information see [Hawkins, 1996, Bamler, 1992, Chang et al., 1992].

Wavenumber The wavenumber algorithm was initially used in radar applications [Cafforio et al., 1991] and is based on a concept from geophysics, namely seismic migration [Stolt, 1978]. In some literature, the wavenumber algorithm is known as the range migration algorithm [Carrara et al., 1995]. The wavenumber algorithm is based on an inverse of the wave equation, with inversions developed by both [Bamler, 1992] and [Soumekh, 1992]. For a more detailed analysis see [Hawkins, 1996]. The wavenumber algorithm operates on the data in the wavenumber domain after across-track pulse-compression has been performed. An important processing step within the wavenumber algorithm is a data remapping known as 'Stolt-mapping' [Stolt, 1978, Hawkins, 1996] requiring accurate interpolation. All the synthetic aperture processing used throughout this thesis is performed using the wavenumber algorithm. The choice of using the wavenumber algorithm has little impact on the interferometric techniques presented in this thesis, primarily chosen because of experience using the wavenumber algorithm within the Acoustics Research Group at the University of Canterbury.

Chirp-scaling The chirp-scaling algorithm combines both pulse-compression and synthetic aperture processing into one algorithm [Cumming et al., 1992, Runge and Bamler, 1992, Raney et al., 1994]. It has one major advantage over the wavenumber algorithm, not requiring the Stolt mapping processing step, thus eliminating the need for computationally expensive interpolation. The algorithm exploits the scaling/shifting properties of a linear frequency modulated signal (LFM chirp), and is thus limited to systems using LFM signals. This may initially seem a disadvantage, but most modern SAS systems do employ a LFM chirp signal, allowing the use of the chirp-scaling algorithm. The efficiency of the chirp scaling algorithm was later improved with the

introduction of the accelerated chirp scaling algorithm [Hawkins, 1996, Hawkins and Gough, 1997], currently the computationally fastest known SAS algorithm.

1.1.5 Motion estimation and correction

A problem for all non-rail based SAS systems is unwanted platform movement. Ideally, the sonar platform moves along the synthetic aperture in a straight line parallel to the scene to be imaged, with a known, constant forward velocity. Under these conditions, the synthetic aperture is evenly sampled, and given adequate along-track sample spacing, a diffraction limited image can be produced. However, the path of the platform through the water will never follow this ideal path, instead there are position and orientation errors along the aperture. The position and orientation of the platform can be estimated from three sources, sensors on the platform, external systems utilising reference points, or from the data itself. There are many sensors which can be incorporated into the platform such as: pressure sensors, accelerometers, Doppler-velocity logs (DVLs), and gyroscopes. These and other sensors are discussed in more detail in Section 5.5. External platform sensors are systems designed to estimate the position and orientation of the sonar platform by triangulation type methods, usually from a network of known reference points. These systems are functionally similar to the Global Positioning System (GPS), often used for positioning above the earths surface[1]. For more discussion on these systems see [Pilbrow, 2007], [Milne, 1983], [Deffenbaugh et al., 1996], and [Vickery, 1998]. The third source of position/orientation estimates is from the sonar data itself, an attractive solution because no extra sensor hardware and data logging is required. This helps reduce the cost and complexity of the platform hardware, although more data processing is normally required. Motion estimation is an active area of current research, with many algorithms based on the concept of displaced phase center antenna (DPCA) [Bellettini and Pinto, 2002], and displaced ping imaging autofocus (DPIA) [Gough and Miller, 2004].

Once an estimate of the towfish motion has been made, the estimate can either be used to distort the motion corrupted data to the ideal straight path, or can be incorporated into the synthetic aperture algorithm. For interferometric applications it is important to apply the same data corrections to all the interferometric elements, else errors will be produced in the resulting bathymetry estimate. Normally this condition is satisfied by estimating the platform motion, then applying this common motion to each of the interferometric receiver elements according to their physical location with respect to the overall platform.

Further complicating the effects of platform motion is the difficulty in accurately measuring all six degrees of freedom independently. Most sensors only measure some

[1]GPS cannot be used directly for underwater positioning since there is no propagation of the GPS signals into the water column. The electromagnetic GPS signal is rapidly attenuated in the conducting sea-water.

of these six degrees of freedom, and sometimes as a combination of the six. Separating the sensor measurements to give an absolute position and orientation of the platform at a given point is difficult, covered in more detail in Section 5.5.

1.1.6 Autofocus

Autofocus algorithms are applied to the imagery after synthetic aperture processing [Gough and Miller, 2003, Fortune, 2005]. As shown in Figure 1.2, the definition of autofocus used here is distinct to motion-correction. Generally, autofocus algorithms are not phase preserving, so the resulting imagery is not suitable for interferometric processing. However, many autofocus algorithms give an estimate of the platform motion, able to be used to correct the data before synthetic aperture processing, providing imagery suitable for subsequent interferometric calculations.

1.1.7 Scene object detection and classification

For many sonar applications, the final data processing stage detects and classifies objects within the scene. Commonly this is refereed to as Computer aided detection/Computer aided classification (CAD/CAC). Bathymetry estimates provide an important source of information for detection and classification of objects, allowing distinctions between objects of similar acoustic signatures. The shape and size of shadow regions within the scene can also be incorporated into the detection and classification of scene objects [Reed et al., 2003, Reed et al., 2004].

1.1.8 Practical SAS image processing considerations

When processing the data collected from a SAS system, care must be taken to ensure numerical errors are not introduced into the data. Often large sources of corruption can be tolerated with intensity based processing, however, the inherent dependence on signal phase of SAS algorithms implies a lower level of numerical error can be tolerated. Even small phase errors can have a pronounced effect on the final imagery from SAS algorithms [Callow, 2003].

Circular effects

As discussed in Section 1.1.4 and shown in Figure 1.3, many SAS algorithms use a combination of the four spatial/frequency domains. Conversion between these domains is performed using a Fourier transform[2], and must be processed with care to avoid circular convolution effects in the image. To avoid such effects, the scene should be surrounded with a region of zero value data points before conversions are performed

[2]Often implemented as a Fast Fourier Transform (FFT) for numerical efficiency.

between the spatial and frequency domains. This process is commonly referred to as 'zero padding'. Often the number of zeros added to the surrounds of the image will be chosen to give an image size with many factors (such as 2^n), to provide maximum numerical efficiency gains of fast implementations of the Fourier transform.

Interpolation schemes

Almost all signal processing requires interpolation steps to re-sample the data. Interpolation must be done with care to ensure the original data is preserved, avoiding the introduction of image distortions and blurring. Interpolation is also often a computationally expensive process, especially when more accurate high order interpolation schemes are used. The choice of interpolation scheme for SAS algorithms is of vital importance for InSAS processing, since any error introduced in the SAS processing stage propagates through to the interferometric processing stages. For InSAS processing algorithms, and important processing step requiring interpolation is resampling the data from the slant-range plane to a common ground-range plane. An analysis of commonly used interpolation schemes is presented in Section 3.7.3.

Area of valid data

When processing datasets from typical strip-map scenes with large extents (many thousand pixels in either/both directions), it is normally easier to decompose the dataset into several smaller blocks. Each of these blocks can then be processed independently, and recombined to give one overall processed scene. This has many advantages computationally since only one of the smaller blocks is required at a given time, also allowing the processing load to be divided over multiple processors if available. However, if the sub-division of the data is performed by simply truncating the dataset into convenient chunks, the resulting image after synthetic aperture processed image is poor around the joins of the recombined data. This image degradation is due to truncation of the image locus for pixels near the edge of each block. Instead, a superior approach is to sub-divide the image into overlapping blocks, such that the overlap distance is greater than the extent of the locus of a point midway between the overlapping portions. After synthetic aperture processing of each of these overlapping blocks is performed, the central portion is retained, discarding half the overlapping region from each side. Concatenation of these central portions yields one image with equal theoretical resolution at all image positions. Since the two outer portions of the image are discarded the synthetic aperture processing for these image positions can also be discarded. However, this is often not possible with range-Doppler and wavenumber domain based synthetic aperture processing algorithms. This overlapping concept is shown in Figure 1.4 with fully overlapping extracted regions. The final stripmap data will also exhibit a small

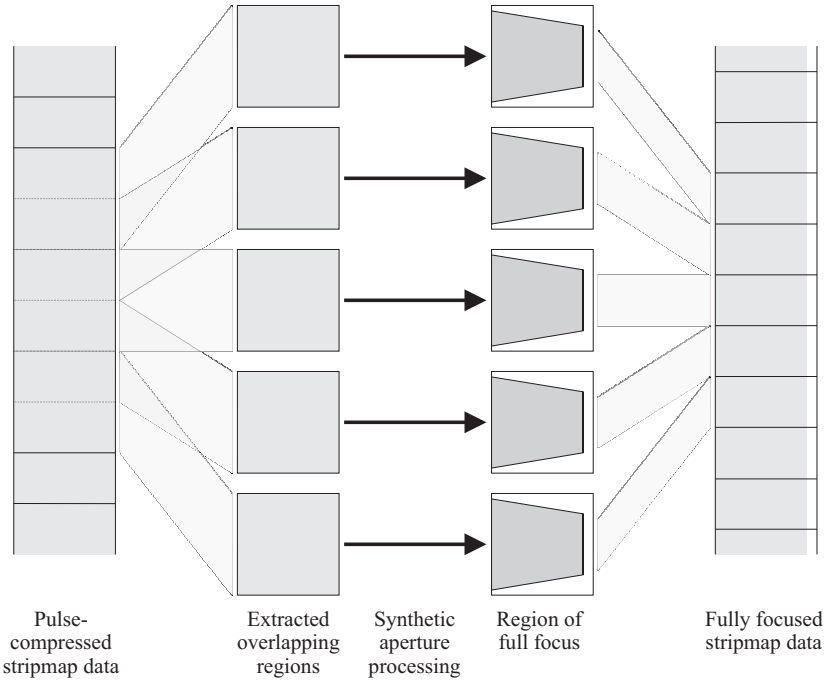


Figure 1.4 For large along-track strip-map data sets, synthetic aperture processing can be performed on overlapping portions of the strip-map data. Each overlapping portion is processed individually as an independent scene, giving a smaller portion of fully focused data. By overlapping the portions sufficiently, the middle (fully focused) region can be extracted from each portion to give the final fully focused strip-map image. Note, there is a small strip of data at far range where the data is not fully focused due to the synthetic aperture locus extending beyond the far edge in the original strip-map data.

strip of poorer resolution at far range where the synthetic aperture locus extends beyond the scene extent; this region should also be considered non-valid, or analysed with an accepted loss of along-track resolution.

It should be noted that the recombination of the processed blocks is not always desirable, since the resulting image can have very large pixel dimensions requiring decimation when displayed on a computer monitor. If the data is presented in a different way (such as a scrolling image on a computer monitor), it may still be desirable to recombine the data into one set. Even if only a small portion of data is to be displayed, this concept of a valid-region should be noted, and a larger area of raw data processed to yield the desired image such that all points in the final image contain the full synthetic aperture.

1.2 ASSUMED BACKGROUND

It is assumed the reader has an understanding of the principles and basic operation of a sonar system. The full details of synthetic aperture processing is not required, nor is an understanding of platform motion estimation and correction. For the interferometric

processing approaches presented in this thesis, it is assumed these processing steps have been performed. A brief introduction to SAS processing was presented earlier in this chapter.

A knowledge of discrete signal processing is required. This includes topics such as: sampling theory, sampled complex baseband signals, Nyquist theory, Fourier theory, spectral filtering, etc. For details of these and other signal processing techniques are covered in many university level textbooks, including:

H. Baher. *Analog & Digital Signal Processing.* John Wiley & Sons Ltd., 1990.

S. Haykin. *Communication Systems.* John Wiley & Sons, Inc., 605 Third Avenue, NY, 3rd edition, 1994.

S. M. Kay. *Fundamentals of statistical signal processing*, volume 1 (estimation theory) 2 (detection theory). Prentice-Hall, Englewood Cliffs, N.J., 1998.

Throughout this thesis there is a slant towards construction and operation of a practical sonar platform. A general understanding of the requirements and operation of an underwater imaging platform would be advantageous to the reader.

1.3 THESIS CONTRIBUTIONS

This thesis contains contributions to the design and processing of a high resolution interferometric synthetic aperture sonar. An analysis of coherence loss for InSAS applications is presented, based upon previous work in InSAR. The design of the sonar interferometric system has been shown to alter the height accuracy performance, in particular, insufficient aperture sampling has been shown to degrade the final estimate. The importance of accurate interpolation is also presented, again causing significant errors in the final height estimate if applied incorrectly.

An analysis of the geometry for a high resolution SAS system is presented. Often the geometry used for the radar equivalent (InSAR) is incorrectly applied to InSAS. The high resolution and slow speed of sound in water cause significant footprint misalignment for the InSAS case, normally insignificant for the InSAR case.

Coherence of InSAS is analysed, comprised of five coherence components. The relative importance of these five coherence components for InSAS is presented, and sonar design strategies presented to reduce the coherence loss of each.

The implication of the along-track sampling spacing on InSAS height estimation is shown. High grating lobe levels can often be tolerated for traditional backscattered SAS processing and imagery, however for InSAS applications, there is a severe reduction in coherence for similar grating lobe levels. The level and effect of the grating lobes can

be reduced by careful selection of the sizes of transmitter and receiver elements and by the along-track spacing of samples along the synthetic aperture.

Belief propagation has been shown to be a useful technique for estimating the height of a seafloor scene from an InSAS dataset. The major advantage of belief propagation is the ability to estimate a seafloor estimate without needing to first correct for the footprint shift between the echo responses. The disadvantage is the height estimate is quantised to discrete heights. Therefore, belief propagation is suited to forming an initial heightmap estimate, from which an initial correction can be applied to the dataset to reduce the coherence loss from footprint mis-alignment.

The design of an interferometric capable SAS system has a large influence on the height estimation performance of the system. In Chapter 5, the selection of the primary parameters for an InSAS system is presented.

The design and construction of a multi-channel data acquisition system for the KiwiSAS-III sonar was integral to this work. This system allowed for sea-trials of an InSAS system, providing imagery for analysis. On overview of the hardware constructed for the KiwiSAS-III system is presented in Appendix B.

1.4 THESIS OUTLINE

Chapter 1 contains introductory material including the concepts of synthetic aperture processing. This processing provides the initial dataset for interferometric processing of further chapters.

Chapter 2 introduces various relevant approaches to remote ground sensing, both on the earth's-surface and underwater. Comparisons are made between radar and sonar systems.

Chapter 3 gives an analysis of coherence as applicable to an InSAS system, divided into the various coherence components of an interferometric sonar system.

Chapter 4 develops a flexible framework based on a maximum likelihood estimator for estimating the time difference of wavefront arrival across an interferometric sonar array. The relationship between coherence and final height estimation error is presented, including the use of multiple independent looks to improve the height error.

Chapter 5 steps through the physical design of an InSAS system. Throughout the process, the implications of various sonar design parameters is presented. Design examples given are largely based on the parameters of the KiwiSAS system. Examples are given from simulated data, with height performance shown for ranges of many design parameters. Sea-trial imagery and results from sea tests of the interferometric capable KiwiSAS–IV sonar system is shown.

Chapter 6 draws conclusions and gives some ideas for areas of further work and investigation.

Appendix A is a reproduction of a conference paper presented on the use of belief propagation as a method of seafloor height estimation. The use of belief propagation is suited to an initial estimate of the seafloor height, then used as an initial correction step for final height estimation.

Appendix B describes the design and construction of the data acquisition hardware to allow the KiwiSAS to operate with interferometric capability.

Chapter 2

INTERFEROMETRIC TERRAIN IMAGING SYSTEMS

Properties of a remotely imaged scene can be categorised into two categories; the overall shape of the scene (hill heights, seafloor bathymetry, etc.), or some property of the surface (colour, density, acoustic reflectivity from sub-wavelength roughness, etc.). Often, it is desirable to produce an image containing data from each of the two groups. For underwater applications, the most common desired combination is acoustic reflectivity and position in 3-dimensional space. Conventional SAS provides high resolution acoustic reflectivity on a 2-dimensional surface, the third dimension (seafloor depth) can be added using an InSAS setup.

2.1 OPTICAL VERSUS COHERENT IMAGING SYSTEMS

When estimating the topography of the surface of the earth, imaging systems fall into two categories: non-coherent and coherent systems. Non-coherent systems are often implemented with a pair or more of cameras configured as a stereo pair. Examples of such systems are as used in the surveying industry; either from an aircraft for large scale imaging, or smaller ground based systems for smaller scale imaging. By estimating the shift in the resulting intensity images, an estimate of the ground topography can be found. These systems often use an ambient energy source, such as the sun, with measurements made from the one way wavefront propagation from target to receiver. Some systems do use a controlled energy source (such as a laser), however, the phase of the reflected wave is not measured so is still considered a non-coherent system.

The second category of imaging systems are coherent systems, where the phase of the incoming wavefront can be inferred. Using the phase difference between two receivers, an estimate of the height can be made. These coherent systems normally use a single controllable energy source, imaging the reflected wavefronts from the scene. The medium through which the imaging is performed dictates the coherent nature of the wavefront. For example, in radar systems where electromagnetic waves are used, the polarisation of the signal can be recorded and used as a further information source. However in underwater acoustic systems, no polarisation of the wavefront is possible.

2.2 MULTI-BEAM ECHO SOUNDERS

Multi-beam echo sounders are usually deployed as a ship hull-mounted, downward-looking system. The system typically consists of two sub-systems, a transmitter and a multi-element receiver. The transmitter transducer has a much greater extent in the along-track direction (fore-aft of the ship) than the across-track direction. Because of these transducer dimensions, the beam-pattern of the transducer is small in extent in the along-track direction, while wide in the across-track direction. The receiver hydrophone has similar dimensions to the transmitter, however, mounted perpendicular to the transmitter. The beam-pattern of this array is similar to the transmitter, except with a narrow width in the across-track direction and a large width in the along-track direction. Unlike the transmitter, the receiver is an array of smaller elements, the echoes of which are recorded individually. By applying an appropriate delay to each of the receiver elements, the look-angle of the receive beam-pattern can be controlled synthetically. When both the transmit and receive beam-patterns are overlaid, the echo returned is restricted to the small area where the two beam-patterns overlap, as illustrated in Figure 2.1. By synthetically steering the receive beam-pattern across a wide range of angles simultaneously, a strip of the seafloor perpendicular to the boat can be imaged. After a small forward movement of the boat, this process can be repeated, imaging another strip across the seafloor.

Seafloor resolution cell size is determined by the beam-width of the composite transmit and receive beam-pattern, and the distance and angle of the sonar system to the seafloor. The resolution cell size can be determined by projecting the beam-pattern onto the seafloor, including the across-track look angle. Thus, the resolution cell size is a minimum (highest resolution) directly below the sonar system, increasing in size as the across-track angle is increased. Typical, multi-beam echo-sounders have an operational swath of 120–150°, with a composite transmit and receive beam-pattern in the range of 0.5–2.0°. For example, the Reson[1] SeaBat 7125 [Reson, 2006] has a swath of 128° consisting of 128 beams of 1° beam-width each for its lower frequency operating band (200 kHz), and/or 256 beams of 0.5° beam-width each for its higher frequency (400 kHz) operating band. The groundplane resolution cell size of this system is shown in Figure 2.2 for a seafloor depth of 20 m. Because the resolution cell size is dependent on the depth of the seafloor, the resolution is poor in deep water, restricting the minimum size of features able to be imaged. For example, at a depth of 100 m, a 1° beam has a footprint of greater than 1.7 m directly below the sonar, increasing to more than 3.5 m for the widest look angle. To decrease the size of the resolution cell, the transmit and receive beamwidths need to be very narrow, requiring very large, impractical, arrays. In shallower water, the resolution cell size is comparable to that of a side-scanning sonar system. The disadvantage over a side-scanning system is the

[1] www.reson.com

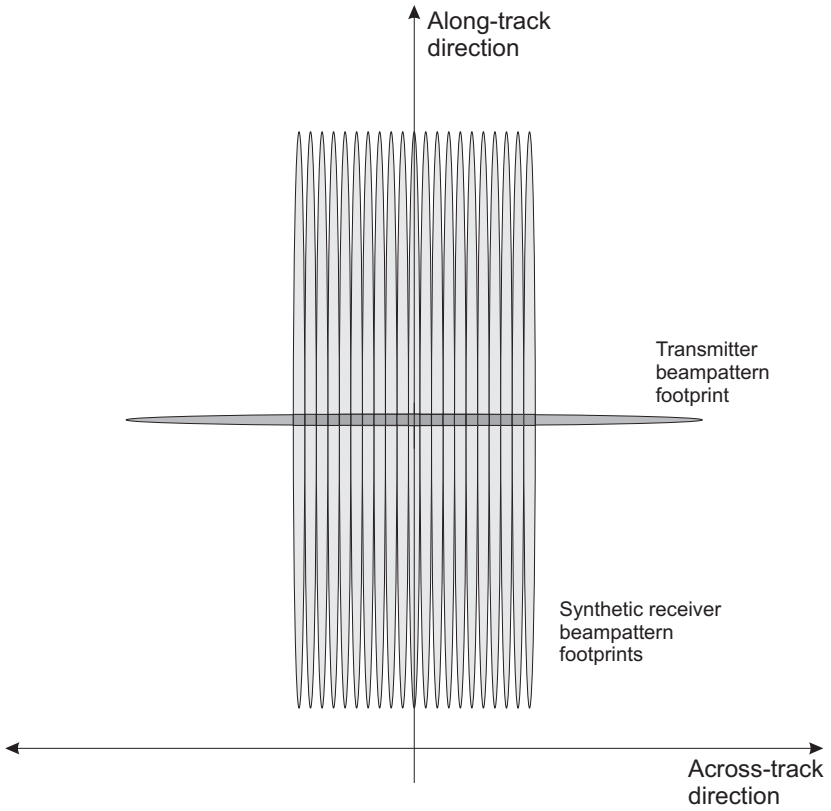


Figure 2.1 Typical beam patterns of a multi-beam echo sounder. The transmit beam-pattern covers a wide portion of the seafloor, with a small along-track extent. The receive beam-pattern has a large along-track extent, with a narrow across-track size. The echo response will only be from the small overlapping area of the two beam-patterns. By synthetically steering the receive beam-pattern to many positions across the seafloor, the point of intersection of the two beam-patterns can be altered, imaging a separate portion of the seafloor.

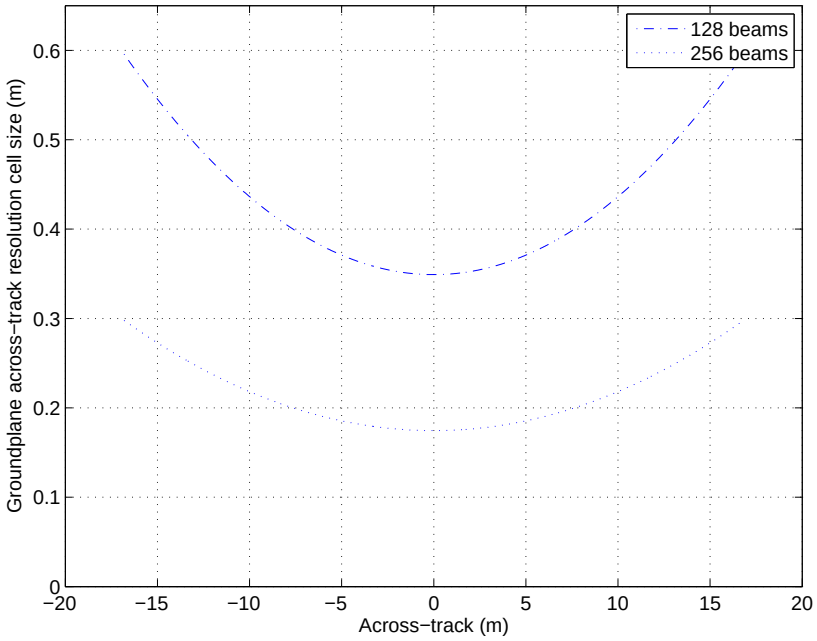


Figure 2.2 Groundplane resolution cell size for a multi-beam echo sounder based on the Reson SeaBat 7125 system with a swath width of 128° and a water depth of 20 m. The swath is synthetically divided into a number of non-overlapping beams of equal angle. Thus, the groundplane across-track resolution cell size increases as the across-track look angle is increased.

small imaging swath due to the down-looking orientation of the multi-beam system.

Since a multi-beam echo-sounder is essentially a downward-looking system, the seafloor height (z direction) measurement accuracy is limited by the resolving power of the acoustic signal used. Typically, these systems use a single frequency, continuous waveform, time-gated to produce a short pulse. By filtering the received echo with a narrowband filter, reflections from objects within the water column and the seafloor can be detected. By assuming a single point of reflection on the seafloor, the time of flight to the seafloor can be determined with high accuracy. The height accuracy is the average height over the groundplane resolution cell, often a large area. Since multi-beam echo sounders are essentially downward looking, objects proud of the seafloor will not cast significant shadows when compared to a side-scan sonar. However, due to the high height accuracy of a multi-beam echo-sounder system, such objects will show significant height differences to the surrounding seafloor. Due to the poor groundplane resolution, the height of small proud objects may be mis-measured because of an averaging effect with the surrounding seafloor within the resolution cell.

2.3 HIGH RESOLUTION SAS INTERFEROMETRY INTRODUCTION

Interferometric synthetic aperture sonar (InSAS) requires an array of two or more hydrophones with vertical separation. Throughout this thesis it is assumed the insonifying signal is from one transmitter, common to all hydrophones. The intensity imagery from each hydrophone from the array is processed separately using phase preserving algorithms, as discussed in Section 1.1. Thus, the starting point for all the InSAS processing presented is using the 'best' imagery from each hydrophone, all processed using the same parameters.

2.3.1 Sidescan InSAS coordinate systems

Absolute positioning of the imagery and bathymetry produced by an InSAS system can be important in situations where a certain point in the scene needs to be returned to at a later date. Such circumstances could arise in object detection surveys where after a target of interest is detected within the InSAS imagery, further exploration is required, either by further sonar surveys or visual verification. Another circumstance where absolute positioning is important is large area surveys, requiring multiple swaths to image the area. Absolute positioning allows the multiple passes to be aligned and registered, ensuring full scene coverage with minimal wasteful overlap between the swaths. Such positioning is of high importance for AUV vehicles where the absolute position must be tracked in real-time to provide positioning data to the real-time navigation system.

During the formation of a bathymetry image from InSAS data, the absolute position of the sonar is of less importance, only a relative positioning is required over the synthetic aperture length. After formation of the imagery and bathymetry within the local coordinate frame has been made, the absolute position can be found by translating the scene into the global coordinate frame. Throughout this thesis the translation from the relative coordinate frame (x,y,z) to a global coordinate frame (e.g., latitude and longitude) is not considered, with only a z axis towfish depth measurement added to the overall bathymetry of the scene.

2.3.2 Sidescan InSAS geometry

A mistake commonly made for broadband InSAS systems is in the definitions of the system geometry. Often the geometry considered is of a single point on the seafloor acoustically illuminated by the transmitter. From this point the acoustic paths are traced back to the receivers, with the difference in path lengths the parameter to be estimated to give the height estimate. This geometry is incorrect for a broadband system with simultaneous sampling of the interferometric receivers. More correctly, the path length of each transmitter/receiver pair is identical (since each receiver is

sampled at the same time offset from signal transmission), thus the point imaged on the seafloor is different for each receiver. The origin of this geometrical error is from the application of narrowband algorithms (for example InSAR) to broadband situations.

The geometry under consideration is a sidescan sonar with a rigid vertical array of identical hydrophones, and a single transmitter. It is also assumed the sonar systems samples all receivers at a fixed rate, with all hydrophone samples occurring at the same time instant (synchronously sampled). Here it is assumed the raw signals from each of the hydrophones has been processed as a synthetic aperture, providing a set of images with range independent resolution in both across- and along-track directions. The total path length from transmitter to a point on the seafloor, then back to each hydrophone can be divided into two parts, the transmit path and the receive path. Although all hydrophones share a common transmitter element, the transmit path cannot be considered common to all hydrophones. Similarly, on the returning receive path, the length is different depending on the hydrophone under consideration. For a given location within the scene the slant-range distances to each receiver can be found from the ground-plane range and depth from a given transmitter/hydrophone pair.

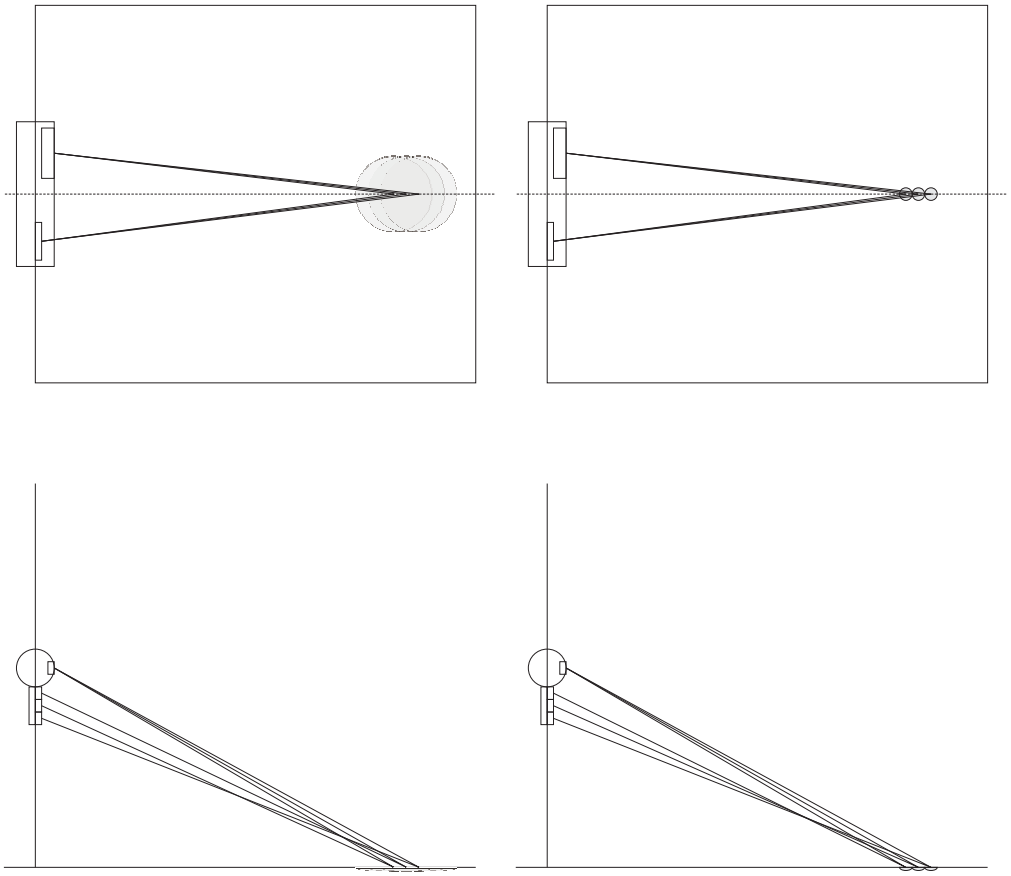
The geometry as defined in Figure 2.3 is based on true distances, rather than a time based variable (time of flight) as is used by the sonar system. Distances cannot be directly measured using a sonar system, instead distances are inferred from the time of flight. To make this conversion from time to distance, the speed of sound, c, through the medium must be known. Unfortunately, the speed of sound in sea-water varies depending on many factors including, but not limited to, temperature, salinity, and pressure [Urick, 1975]. Hence, a truly constant speed of sound cannot be defined, however, throughout this thesis it is assumed constant,

$$c = 1500\,\mathrm{m/s}. \tag{2.1}$$

For a practical sonar system, the imaging geometry is the reverse of that shown in Figure 2.3. Instead of considering a point within the scene and calculating the distance from the transmitter and back to the hydrophones, it is more correct to consider the point within the scene at a given time (or equivalently distance) offset. Thus, a point within the scene will appear at a different position in each of the hydrophone signals. The position is further complicated by the sampling nature of a practical sonar system. However, this complication can often be ignored provided the signal is sampled at a sufficiently high rate to obey Nyquist sampling theory, and is considered at all points throughout the signal processing chain.

2.3.3 Narrowband time delay estimation

Consider an interferometric system with two ($N_h = 2$) receiver elements configured as a single interferometric pair, imaging with a narrowband signal centred at f_c. At some



(a) Narrowband InSAS geometry; birds-eye view top, end-on view bottom.

(b) Broadband InSAS geometry; birds-eye view top, end-on view bottom.

Figure 2.3 InSAS geometry for, (a) narrowband case, and (b) broadband case. In the narrowband case resolution is poor giving large, overlapping, resolution cells for each of the transmitter receiver pairs. In the broadband case the resolution cells can be sufficiently small to cause separation of each of the transmitter receiver pairs for the same system geometry as the narrowband case. The large overlap of the resolution cells in the narrowband case means each of the transmitter receiver pairs of the interferometric array effectively image the same portion of seafloor. The broadband case has little or no overlapping portion between the interferometric array so direct interferometric measurements become meaningless.

point of interest within the scene, the signals from the two receivers can be modeled,

$$d_1 = a_1 m_1 + n_1, \tag{2.2}$$

$$d_2 = a_2 m_2 \exp\left(j2\pi b f_c \tau\right) + n_2, \tag{2.3}$$

for hydrophones $h \in (1, 2)$, where a_h is the (complex) signal amplitude, m_h models the multiplicative noise of speckle, and n_h models all additive noise. The time delay parameter to be estimated due to hydrophone separation is τ, and b is the effective baseline between the two receivers. The interferometric problem is to estimate τ from the data vector $\mathbf{d} = [d_1, d_2]^T$. Since the signal is narrowband, the across-track resolution of the system is poor. In this case the difference in path length from a point of interest to each of the receivers is smaller than the across-track resolution size. As a result, the effect of τ is phase-only, so the interferometric problem is equivalent to estimating the phase difference of data vector $\mathbf{d}$.

2.3.4 Narrowband interferogram

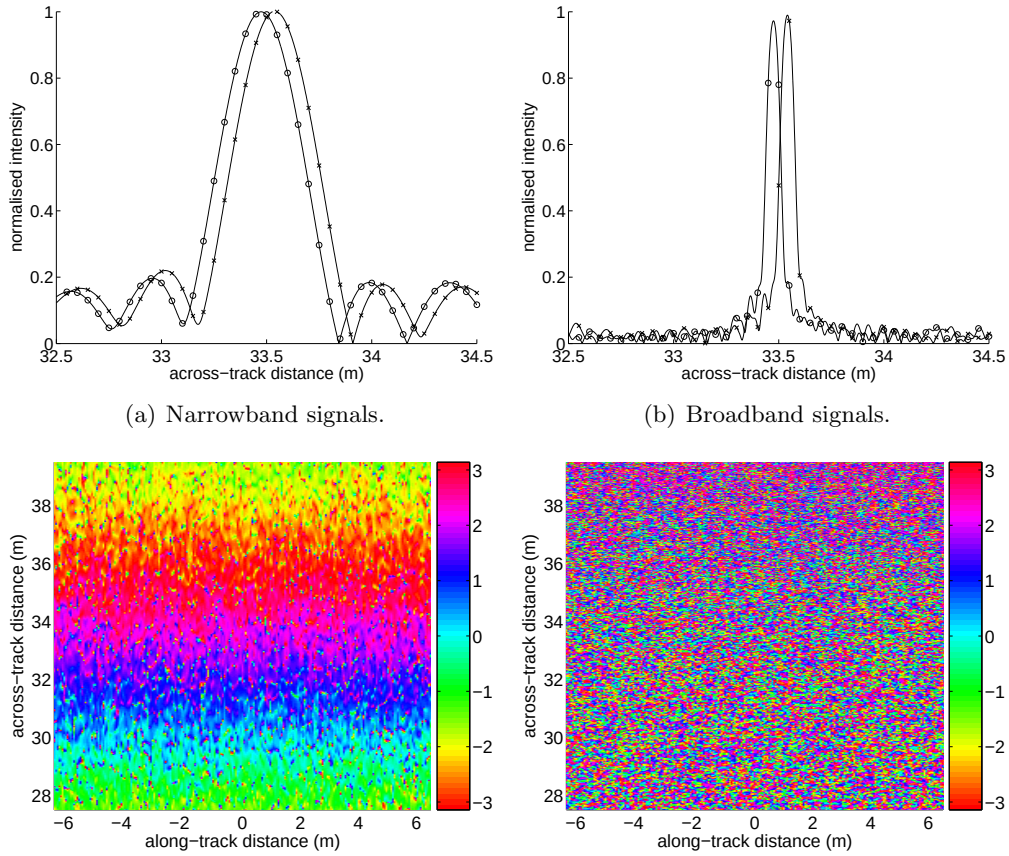
An interferogram can be formed between the two signals of (2.3), formed by a pixel-by-pixel phase difference,

$$\hat{\phi} = \angle\left(d_1 d_2^*\right). \tag{2.4}$$

The result of forming an interferogram on narrowband data can be seen in Figure 2.4.

Phase unwrapping

For phase-based calculations in interferometry applications, it is necessary to estimate the absolute phase difference across the baseline. Unfortunately, estimation of the phase difference across the interferometric baseline will normally be modulo-2π. Calculating the correct number of 2π 'unwraps' to apply to the modulo-2π data is non-trivial, especially when considered across a two-dimensional dataset. The concept of 'branch-cuts' was introduced by [Goldstein et al., 1988] as a two-dimensional phase unwrapping algorithm, specifically applied to InSAR. The technique introduced by Goldstein has difficulty in correctly unwrapping the phase fully in certain data sets, especially when the scene has abrupt phase change boundaries spanning the entire scene. Many other algorithms have since been posed as improvements [Xu et al., 1994, Massonnet et al., 1996, Lombardini, 1996, Lombardini and Lombardo, 1998, Corsini et al., 1997, Lombardini, 1998, Kim and Griffiths, 1999, Chen and Zebker, 2000, Chen and Zebker, 2002]. A good summary of many of the modern algorithms is a book by [Ghiglia and Pritt, 1998], including implementation issues and execution speed. Many of these two-dimensional phase unwrapping algorithms have also been applied to InSAS [Banks et al., 2000, Banks, 2002], however, the fundamental differences between the original narrowband applications (InSAR) and the broadband InSAS application are often not



(a) Narrowband signals.

(b) Broadband signals.

(c) Interferogram of narrowband signals.

(d) Interferogram of broadband signals.

Figure 2.4 Simulated point on a rough flat seafloor, received by a pair of hydrophones with 0.3 m vertical separation. Centre frequency $f_c = 30\,\text{kHz}$, bandwidth $BW = 2\,\text{kHz}$ for (a) the narrowband case, and $BW = 20\,\text{kHz}$ for (b) broadband case. Note the distance offset between the signals is equal in both cases, but the loss in overlap is greater in the broadband case. The shift has a marked impact on interferograms formed on a per-pixel basis. For the narrowband case (c), the interferogram phase can be seen to be slowly changing across the seafloor patch. For the broadband case (d), the mis-registration causes the interferogram phase to be random.

considered. The application of two-dimensional phase unwrapping techniques cannot be applied directly to the broadband case due to the mis-registration between the images. Although corrections could be made to register the images to within one pixel as required, determining the correction required is often based on time-domain correlation algorithms, valid only when a narrowband subset of the original broadband data is formed.

2.3.5 Broadband time delay estimation

If the imaging signal is broadband, algorithms based on narrowband assumptions will often not work correctly. The overall interferometric problem is still the same as the narrowband case, estimating the time difference τ between the vertically separated receivers. However, with a broadband SAS signal the across-track resolution cell size is smaller in extent than the path difference. This causes a significant time shift between the receiver signals, hence the narrowband approximation of a phase only difference cannot be applied. Forming an interferogram from broadband SAS data results in a phase-map of random phase, as seen in Figure 2.4.

Time-domain cross correlation

A time domain cross correlation can be used to estimate the time difference between the two receiver signals for a broadband system. After forming the cross-correlation, the peak is found, the shift from zero lag giving the time-difference estimate. The peak shift must be found to sub-pixel accuracy, implying the need for a combination of highly sampled data and/or sub-pixel interpolation. Constraints on the maximum expected time shift between the receivers can be calculated from the receiver separation and loose scene bounds, limiting the lags over which the peak search must be computed. The cross-correlation is formed from a small section of data from each receiver, centred around some point of interest. The size of this section of data determines the ground-plane resolution over which the bathymetric estimate is made. For example, during the bathymetric processing from the SAMI system [Châtillon et al., 1999], the correlation length is stated as in the order of three resolution cells, giving a single height estimate for a portion of scene three times the across-track resolution. The SAMI system has a stated resolution of 0.3 m in across-track, and 1 m in along-track, hence with a correlation length of three resolution cells, the bathymetric estimate will be over a roughly square 1 m $\times$ 1 m area. Performing the bathymetric height estimate over a set of across-track resolution cells is equivalent to reducing the imaging bandwidth of the system. This reduction in bandwidth effectively converts the system from broadband to a narrowband system. So although the cross-correlation is performed on full bandwidth, high resolution data, the bathymetric height estimate produced is at a much lower ground resolution.

2.4 COMMON SAR/SAS RESEARCH PLATFORMS

Listed here are a selection of SAR and SAS platforms for comparison. Some of these systems also have interferometric capabilities.

2.4.1 European Remote Sensing Satellites, ERS-1 and ERS-2

ERS-1 is a European Space Agency satellite devoted to remote sensing from a polar orbit. ERS-1 was launched on 17 July 1991 by an Ariane 4 launcher from Kourou, French Guiana. The mean altitude of flight is 785 km in a sun-synchronous, near polar, near circular orbit, at an inclination of 98.5°. Primarily, the orbit has a 35-day repeat cycle, although other orbits of 3-day and 168-day repeat cycles have been performed. The primary sensor on-board of interest is the active microwave instrument (AMI), combining the functions of a synthetic aperture radar (SAR) and a wind scatterometer (WNS), the operation of which are mutually exclusive. Since there is no on-board data recording capability, AMI data can only be obtained if a ground station is within view. ERS-2 was launched in April 1995 with a similar synthetic aperture radar imaging platform. Parameters of the ERS systems listed in Table 2.2, obtained from [Gatelli et al., 1994].

2.4.2 Airborne VHF SAR, CARABAS

CARABAS (Coherent All RAdio BAnd Sensing) is an ultra-wideband, widebeam SAR operating in the upper HF and lower VHF range. The system was designed and built by the FOA (National Defense Research Establishment, Linköping, Sweden) over the period 1985-1992. Its first test flights were performed during 1992, with the first major SAR trial during October 1992. A small business jet, a Rockwell Sabreliner, is used as the platform for the SAR sensor. The antenna system consists of two dipole aerials attached to the tail section of the aircraft. These antennas are flexible, hollow tubes inflated via an open forward end, with dipole elements of various lengths sewn into the sleeve. These elements are then tuned using analog filters for each frequency of operation. The total length is 5.5 m and a diameter of 0.3 m. The electronics of the system are housed within the aircraft, recording the raw data onto tape for further post-processing on the ground. Parameters of the CARABAS system are listed in Table 2.2, obtained from [Gustavsson et al., 1993] and [Ulander and Frölind, 1998].

2.4.3 University of Canterbury sonar, KiwiSAS

The KiwiSAS system is an ongoing research system, originally designed as a SAS system, developed by the Acoustics Research Group at the University of Canterbury, New Zealand[Hawkins, 1996, Hayes and Gough, 1999, Hayes et al., 2001, Hayes et al.,

2005, Barclay et al., 2005]. Recently this platform has been extended to an InSAS system, initial the KiwiSAS-III. The first iteration of InSAS configuration is covered in more detail in Appendix B and [Barclay et al., 2002a]. After initial proof of concept trials were performed, the system was completely rebuilt into the KiwiSAS-IV system, based on an on-board custom multi-channel data acquisition system. Details of the new system can be found in [Hayes et al., 2002] and [Hayes, 2003].

The KiwiSAS sonar system is designed to provide high resolution imagery of the seafloor in a shallow water environment (water depth less than 20 m), such as inner harbours. The system uses two simultaneously pulsed, linearly chirped FM signals each of 20 kHz bandwidth, each with a centre frequency of 30 kHz and 100 kHz. The returned echoes are received using three vertically separated hydrophone arrays, stored on hard disks for post-processing. The sonar towfish body is slightly positively buoyant, operating at approximately mid-height of the water column by a depressor chain. The towfish is nosed towed from a surface vehicle, and has a flat front face and trailing fins to aid stability. The full parameters of the current InSAS configuration is shown in Table 2.1.

2.4.4 Synthetic Aperture Mapping and Imaging (SAMI)

The Synthetic Aperture Mapping and Imaging (SAMI) sonar system was a collaborative project of five European partners. The system was designed as a wide-band side-scan sonar system to produce high-resolution imagery and bathymetry. The sonar system consisted of a towed towfish body containing the sonar arrays and associated electronics, connected to the tow vehicle by a cable[Lawlor et al., 1996, Châtillon et al., 1999]. On-board the two vehicle was the navigation and control units, sonar processing, storage, and display units. The towfish housed a 2×2 array of transducers, each of 1 m length, 0.26 m in height. One of the four elements was used as a transmitter, all four used as receivers. The acoustic signal covered the frequency range 5-10 kHz, with a stated signal strength of 218 dB ref 1 μPa at 1 m. The arrays could be combined in several combinations to form two signals, each fed to matched-filtering (pulse compression) and synthetic aperture processing units. The matched-filtering was performed using an array of Motorola DSP devices, the synthetic aperture processing performed using an array of transputers. Each transputer focused a different range portion of the scene, with load balancing achieved by splitting the data according to synthetic aperture length requirements. The stated across-track resolution is 0.3 m, with a theoretical along-track resolution after SAS processing of 0.5 m or 1.0 m depending on the receiver summation configuration[2]. The towfish body had a depth rating of up to 250 m, with

[2]The stated resolution is greater (poorer) than the theoretical across-track resolution of $\frac{c}{2B} = \frac{1500}{2\times5000} = 0.15$ m. One possible explanation is the frequency range of 5-10 kHz is stated at -6 dB, hence the effective bandwidth is less than 5 kHz. It may also allow some spreading from non-ideal matched-filtering.

Dimensions:	
Overall length	1.7 m
Weight	60 kg
Nominal Origin (x,y,z)	0, -0.850, 0 m
Transmitter:	
Position (x,y,z)	0.095, -0.039, 0.025 m
Dimension (y,z)	0.336, 0.084 m
Orientation	12° down
Signal:	
Centre Frequency	30 kHz and 100 kHz
Bandwidth	20 kHz for each band
Type	Tapered linear FM up-chirp
Duration	12.5 ms
Repetition Period	66.66 ms
Receiver array:	
Position - top (x,y,z)	0.0140, -0.6980, -0.2346 m
Position - mid (x,y,z)	0.0140, -0.6980, -0.3120 m
Position - bot (x,y,z)	0.0140, -0.6980, -0.3894 m
Dimension of each (y,z)	0.2322, 0.0774 m
Orientation	0°
Echo:	
Sample Rate	312.5 kHz
Resolution	16 bit
Storage	250 GB
Operation:	
Depth	Typically 10 m
Forward Velocity	1.5 – 3 kn

Table 2.1 Parameters of the KiwiSAS-IV system.

Parameter	Symbol	Units	ERS-1	CARABAS	KiwiSAS-III
Wave propagation speed	c	m/s	3×10^8	3×10^8	1500
Centre frequency	f_c	Hz	5.3×10^9	40×10^6	$30/100 \times 10^3$
Wavelength	λ	m	0.0567	7.5	0.05/0.015
Bandwidth	BW	Hz	16×10^6	20×10^6	20×10^3
Quality factor	Q	-	331	2	1.5/5
Imaging Distance	H	m	780×10^3	7500	15-200
Along-track Speed	v	m/s		100	1
Resolution		m	9×9	8×8	0.05×0.15
Critical Baseline	B_c	m	1100		
Altitude		m	785×10^3	$1500 - 4500$	5

Table 2.2 Comparison of common InSAR and InSAS systems.

a operational swath of 100 –2500 m.

2.5 COMPARISON OF PARAMETERS

When comparing InSAR systems with InSAS systems, comparisons fall into two major groups: the geometry of the imaging system, and the operating parameters of the systems. The main parameters of some of the common InSAR and InSAS systems is shown in Table 2.2.

2.5.1 InSAR and InSAS geometry comparison

One major difference between the two imaging systems is the look angle of the system, relative to the terrain being imaged. Typically, InSAR systems operate at a greater angle below zenith, with a small difference in angle of arrival across the imaging swath. Due to the often shallow water environment of InSAS systems, the imaging is typically only slightly below horizontal for the furtherest imaging point, with a large difference of angle of arrival due to the desirably large imaging swath.

Multipath from the sea surface has shown to be a major problem in InSAS systems not seen in InSAR systems since there is no analogue of the sea surface. However, there is much current research in InSAR dealing with volumetric scattering from areas of the earths surface such as forests. Since there is some sound penetration into some seafloor substrates, this may be an area of future research for InSAS systems.

2.5.2 InSAR and InSAS operating parameters comparison

The primary difference between InSAR and InSAS systems is the quality factor, or Q, of the system. In the case of satellite based systems such as the ERS platforms, the imaging band is centred in the gigahertz range with a bandwidth typically in the tens

of megahertz, giving an overall Q in the hundreds. For InSAS platforms the resulting Q is much lower, often in the 2–5 range. This has the serious consequence of often making narrowband approximations invalid for InSAS systems. This large bandwidth can, however, be filtered into several smaller narrower bands, to which narrowband approximations can be applied. One InSAR system operating with a similar Q to that of InSAS is the aircraft based CARABAS system.

One other important difference between InSAR and InSAS systems is the length of the interferometric baseline when expressed as wavelengths. Typical baselines for satellite InSAR systems are often in the region of hundreds of meters, translating to several thousand wavelengths. Conversely, InSAS systems typically have baselines of less than half a metre, translating to just a few wavelengths. This difference drastically changes the phase unwrapping techniques required for the two platforms, and can often be avoided in InSAS systems.

The consequences of these differences in parameters between InSAR and InSAS systems is the dominant limiting factors of bathymetry reconstruction differ. With a narrowband InSAR system, effects such as baseline decorrelation become significant, whereas in an InSAS system this has little effect. However, the converse is true for other some factors such as footprint shift. The details of baseline decorrelation are covered in Section 3.5, and footprint shift in Section 3.4.

Chapter 3

COHERENCE OF INTERFEROMETRIC SYNTHETIC APERTURE SONAR SIGNALS

Interferometric systems rely on estimating the time difference in arrival across the interferometric receiver array. The most direct method to estimate the time difference between the interferometric signals is to find the location of the peak in the cross-covariance (by assuming the signals are zero mean signals a cross-correlation can be used). The peak occurs at the time shift equal to the true interferometric delay. With broadband signals typical of InSAS systems, the formation of a time domain cross-correlation may yield maximums at time lags differing from the true interferometric time difference. The amplitude of the peak in the cross-correlation gives a measure of the confidence in the estimate. Several techniques for estimating the correct peak of the cross-correlation are presented in Chapter 4.

In the case of narrowband signals the time-delay estimation process is simplified. The time-difference between the signals is small compared to the duration of the signal. This allows the amplitude of the true peak in the full cross-correlation to be approximated by the cross-correlation at zero lag. The correlation coefficient at zero lag is commonly referred to as coherence.

Complex coherence is classically defined in [Born and Wolf, 1999] as

$$\Lambda = \frac{E\left(d_1 \cdot d_2^*\right)}{\sqrt{E\left(|d_1|^2\right) E\left(|d_2|^2\right)}}, \tag{3.1}$$

where d_1, d_2 are the two signals, and $E(\cdot)$ is the expectation operator. The degree of coherence is the coherence magnitude, $D = |\Lambda|$, and the effective phase difference of the two signals is the argument of Λ. Normally, Λ cannot be calculated for InSAS, since the expectation operation must be performed over all realizations, implying the need for an impractically large number of passes of the scene. Instead, an estimate of Λ is made by replacing the ensemble averages with spatial averages, giving the complex coherence estimator γ. The spatial averaging is performed over a small portion of the scene surrounding the point of interest, over which it is assumed the scatters are independent with identical statistical properties. For this approximation to be valid for

interferometric SAS systems, a phase difference correction term ($\Delta\phi$) must be applied at each pixel location[1] to correct for the change in phase difference due to the change in geometry from the point of interest. Note, the correction term is not the interferometric phase difference, rather a variation of the interferometric phase difference across the region. For a small rectangular region of size N_m in across-track and N_n in along-track (overall size $N_c = N_m \times N_n$), the complex coherence estimator becomes,

$$\gamma = \frac{\sum_{n=1}^{N_n}\sum_{m=1}^{N_m} d_1[m,n]\, d_2^*[m,n] \exp\left(-j\Delta\phi[m,n]\right)}{\sqrt{\sum_{n=1}^{N_n}\sum_{m=1}^{N_m}|d_1[m,n]|^2 \sum_{n=1}^{N_n}\sum_{m=1}^{N_m}|d_2[m,n]|^2}}. \tag{3.2}$$

Again, the estimate of degree of coherence is $d = |\gamma|$, and the interferometric phase given by the argument of γ. As has been shown by [Touzi and Lopes, 1996], the sample coherence magnitude d is biased towards higher values, especially for areas of low coherence and/or small spatial average size (small N_c). The bias can be decreased by increasing the area over which the averaging is performed, at the expense of resolution in the coherence estimate across the scene. Care must be taken when using large windows in scene portions with rapidly changing coherence, typical around strong reflections from complex object shapes. When using large windows, the single coherence estimate made over the region will be an 'average' coherence, and may hide some of the subtle features of the scene.

3.1 INTERFEROMETRIC SAS COHERENCE ESTIMATION

When estimating the coherence of and interferometric pair of signals using (3.2), the phase correction term $\Delta\phi$ must be known over the $N_m \times N_n$ scene area. This implies accurate prior knowledge of the scene topography, a parameter unknown at this stage of the processing [Fowler, 1993]. Instead, the coherence magnitude d can be estimated with an amplitude only estimator, such as the estimator appearing in [Guarnieri and Prati, 1997], dubbed a "Quick and Dirty" coherence estimator. From [Guarnieri and Prati, 1997], the sample correlation coefficient is estimated as,

$$\hat{d} = \begin{cases} \sqrt{2\hat{\rho}-1} & \hat{\rho} > \frac{1}{2} \\ 0 & \hat{\rho} \leq \frac{1}{2} \end{cases}, \tag{3.3}$$

where

$$\hat{\rho} = \frac{\sum_{n=1}^{N_n}\sum_{m=1}^{N_m}|d_1[m,n]|^2 \cdot |d_2[m,n]|^2}{\sqrt{\sum_{n=1}^{N_n}\sum_{m=1}^{N_m}|d_1[m,n]|^4 \sum_{n=1}^{N_n}\sum_{m=1}^{N_m}|d_2[m,n]|^4}}. \tag{3.4}$$

[1]For InSAS systems the phase difference correction term will often only vary in the across-track direction.

In a later publication by [Touzi et al., 1999], the Siegert relationship was used to develop an alternative estimator of d,

$$d^{\text{sig}} = \sqrt{r}, \tag{3.5}$$

where r is equivalent to $\hat{\rho}$ of (3.4) with terms added to remove any non-zero mean,

$$r = \frac{\sum_{c=1}^{N_c} |d_1[m,n]|^2 \cdot |d_2[m,n]|^2 - \sum_{c=1}^{N_c} |d_1[m,n]|^2 \sum_{c=1}^{N_c} |d_2[m,n]|^2}{\sqrt{\sum_{c=1}^{N_c} \left(|d_1[m,n]|^4 - \sum_{c=1}^{N_c} |d_1[m,n]|^4 \right) \sum_{c=1}^{N_c} \left(|d_2[m,n]|^4 - \sum_{c=1}^{N_c} |d_2[m,n]|^4 \right)}}. \tag{3.6}$$

Note the double summation over N_n and N_m in (3.4) has been replaced with a single summation over the area of N_c for compactness.

3.2 COHERENCE LOSS

In in ideal interferometric system, the coherence between the interferometric receiver pairs is unity. However, this is never the case for a practical system, with the coherence often much below the ideal unity. The overall interferometric coherence is the combination of many factors, each dubbed a 'coherence component'. For InSAS, the five main factors are coherence components due to additive acoustic noise, footprint misalignment, baseline decorrelation, temporal decorrelation, and processing noise. These coherence components are designated γ_n,γ_m,γ_b,γ_t, and γ_p respectively. Assuming all the coherence components are independent, the overall coherence is a product of all the coherence components,

$$\gamma = \gamma_n \gamma_m \gamma_b \gamma_t \gamma_p. \tag{3.7}$$

The dominance of each of the coherence components differs depending on the interferometric system. For example, for InSAR applications the dominant coherence component is that of baseline decorrelation (γ_b), however for InSAS this affect is negligible [Touzi et al., 1999]. Conversely, the effect of footprint misalignment (γ_m) is dominant for InSAS, normally ignored for InSAR systems. Each of the five coherence components is explained and analysed in Sections 3.3-3.7, as applicable to the InSAS case. An example is also given in each coherence component section, based on the KiwiSAS system parameters as given in Table 2.1.

3.3 ACOUSTIC NOISE

The coherence loss as a result of additive Gaussian noise is related to the SNR of the noise [Zebker and Villasenor, 1992, Just and Bamler, 1994]. For an interferometric pair of receivers, each with a signal to noise ratio SNR $_h$, the acoustic noise coherence

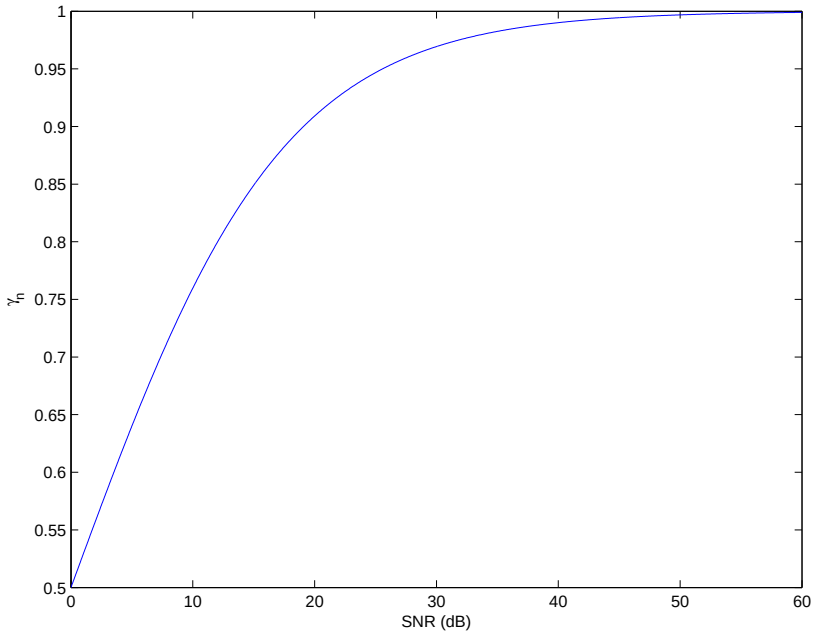


Figure 3.1 The signal coherence is reduced as the SNR is lowered. Significant losses in coherence can result from poor SNR.

component is,

$$\gamma_n = \frac{1}{\sqrt{1+\text{SNR}_1^{-1}}\sqrt{1+\text{SNR}_2^{-1}}}. \tag{3.8}$$

If all the interferometric channels have independent noise with equal SNR, (3.8) reduces to

$$\gamma_n = \frac{1}{1+\text{SNR}^{-1}} \tag{3.9}$$

$$= \frac{\text{SNR}}{\text{SNR}+1}, \tag{3.10}$$

as shown in Figure 3.1. Thus, for an acoustic SNR after SAS processing of 100 (40 dB), the acoustic coherence component $\gamma_n = 0.9901$. Acoustic noise as a source of coherence loss is of particular importance since it cannot be removed by post-processing, and will often impose an upper limit on the overall coherence of an interferometric system.

The relationship between noise and coherence of (3.10) can also be used for analysis of other coherence loss components. If the signal noise/distortion can be expressed as an effective signal to noise ratio, a corresponding coherence component can then be derived using (3.10). An example of this is in Section 3.4 where footprint misalignment can be expressed as an effective SNR. Conversely, if the coherence between two signals is known an effective SNR can be derived, although the mechanism of noise will be

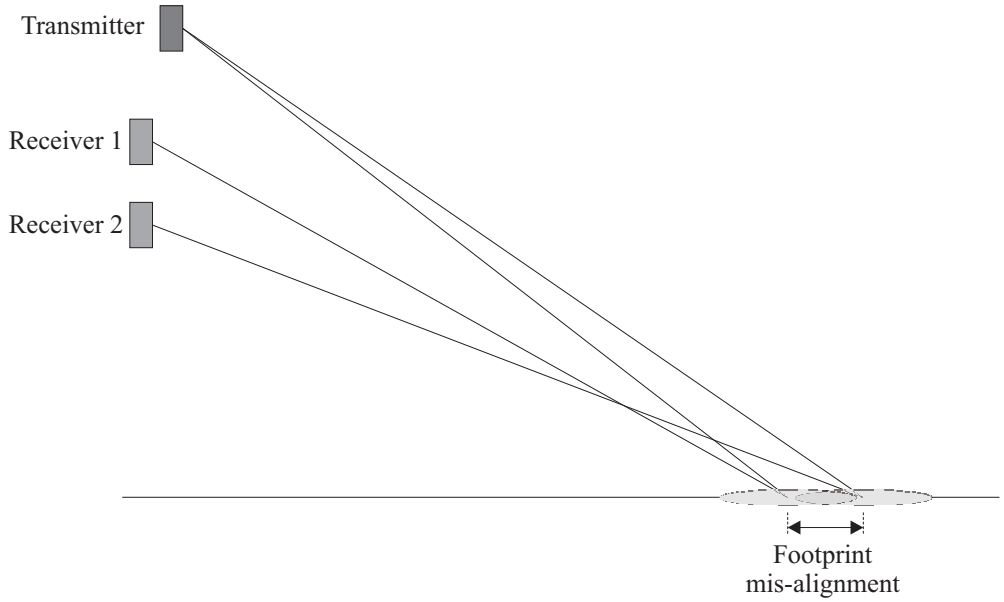


Figure 3.2 Each transmitter/receiver pair has a different footprint position on the seafloor for a fixed acoustic path time/distance. Shown is the case of one transmitter and two receivers with partial overlap of footprints on the seafloor.

unknown. Inverting (3.10) gives

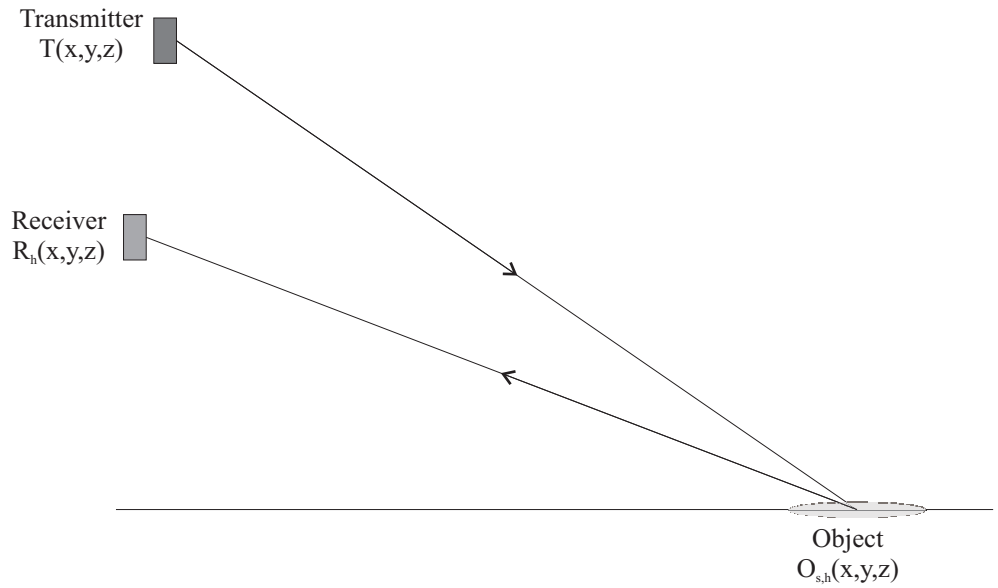
$$\underset{\text{eff}}{\text{SNR}} = \frac{\gamma}{1-\gamma}. \tag{3.11}$$

3.4 FOOTPRINT MISALIGNMENT

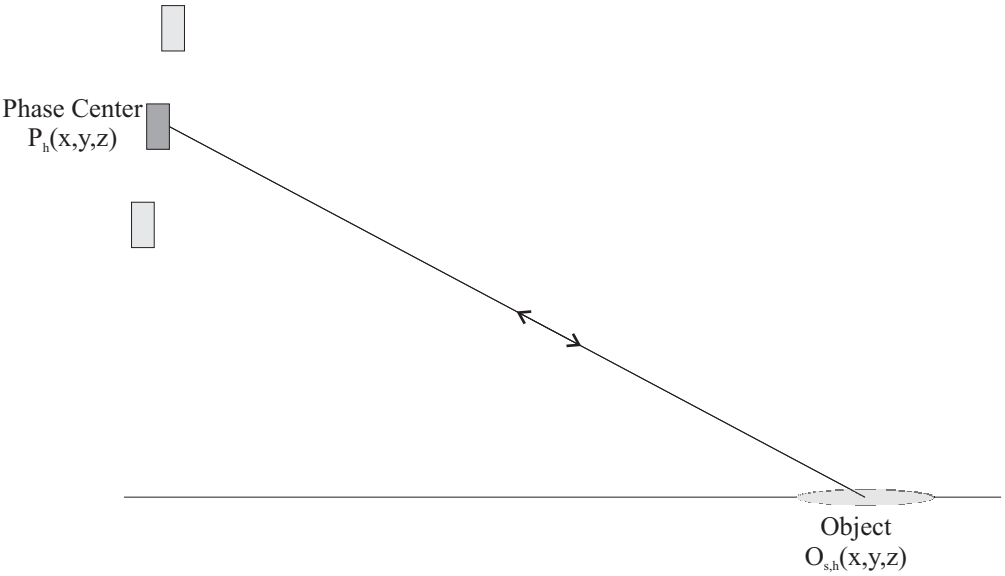
The acoustic path for a side-scanning InSAS system consists of the path from the transmitter (T) to the point of interest, and the return path to a given receiver (R_h). For a time-sampled system, the total time of the acoustic transmit and receive paths is determined by the sampling rate of the system, and the sample count from the time of transmission. For the receivers of an interferometric pair, pixels with the same image location (pixel count) have the same total transmission and reception time. The implication is the position of reflection from the seafloor surface will differ for each of the receivers, due to the height separation of the interferometric baseline. For a given time of flight and known overall scene topography, a single point intersects the seafloor[2] for each of the interferometric transmitter/receiver pairs. Extending around this central point of common length is the across-track footprint. An example using a elliptical footprint shape is shown in Figure 3.2 for a given, single, flight time/distance.

Using the geometry and notation as defined in Figure 3.3(a), the point of reflection

[2]More than one intersection with the seafloor can occur for certain topographies. See Section 5.8 for detail of when layover occurs.



(a) Footprint geometry



(b) Footprint geometry from phase centre

Figure 3.3 Geometry for each transmitter receiver pair of a side-scan InSAS system. Transmitter located at $T(x, y, z)$, and receivers at $R_h(x, y, z)$ with N_h receivers. At a given time delay τ from time of transmission, the total path length is given by r_τ. The intersection of the overall acoustic path occurs a position of $O_{h,\tau}(x, y, z)$. The geometry is simplified by using the phase centre approximation as shown in (b), where the acoustic path is approximated as twice the distance from the phase centre location $P_h(x, y, z)$.

on the seafloor for a given total path length can be calculated from the position of the transmitter, receiver, and seafloor depth. Considering just the across-track and depth dimensions (y=0), from trigonometry the total path length is,

$$\begin{aligned} l(t) &= \sqrt{(O_{h,t}(x) - T(x))^2 + (O_{h,t}(z) - T(z))^2} \\ &\quad + \sqrt{(O_{h,t}(x) - R_h(x))^2 + (O_{h,t}(z) - R_h(z))^2}, \end{aligned} \tag{3.12}$$

where T is the transmitter position, R_h is the position of receiver h, and $O_{h,t}$ is the position of scene intersection for a time delay from transmission of t. Solving for the object across-track position yields a 4^{th} order polynomial. A simplified expression can be found by using the phase centre approximation, as shown in Figure 3.3(b). Instead of calculating the distance from the transmitter to object and back to a given receiver, a two-way path is traced from the phase centre to the object. Using the phase centre approximation,

$$l(t) \approx 2\sqrt{(O_{h,t}(x) - P_h(x))^2 + (O_{h,t}(z) - P_h(z))^2}, \tag{3.13}$$

where P_h is the phase centre position for transmitter/receiver pair h. Solving for the across-track object position gives,

$$O_{h,t}(x) = P_h(x) + \sqrt{\frac{r(t)^2}{4} - (O_{h,t}(z) - P_h(z))^2}. \tag{3.14}$$

A second solution also exists in the negative across-track direction, backwards from the direction of the transmitter and receiver. This solution can be ignored due to geometric prior knowledge.

To calculate the footprint mis-alignment of the interferometric pair of receivers, the across-track object position is calculated for each receiver using (3.14), the difference giving the mis-alignment,

$$\varphi_m = |O_{1,t} - O_{2,t}|. \tag{3.15}$$

By considering the footprint misalignment as a source of Gaussian noise, an equivalent signal to noise ratio can be found. The overlapping portions of the two footprints are considered as 'signal', the non-overlapping portions 'noise'. This is an approximate measure, valid when:

- the non-overlapping portion is completely uncorrelated to the overlapping portion, while the overlapping portions are completely correlated.
- the signal spectrum is uniform over the entire footprint.
- the two footprints, and therefore the two non-overlapping portions are identical in duration.

For a rectangular resolution cell of across-track size φ_r satisfying these assumptions, the equivalent SNR is,

$$\underset{m}{\text{SNR}} = \frac{\varphi_r - \varphi_m}{\varphi_m} \tag{3.16}$$

From the equivalent SNR , the coherence component γ_m can be calculated using (3.10),

$$\gamma_{m(\text{rect})} = \frac{1}{1 + \text{SNR}_{\,m}^{-1}} \tag{3.17}$$

$$= \frac{\varphi_r - \varphi_m}{\varphi_r}. \tag{3.18}$$

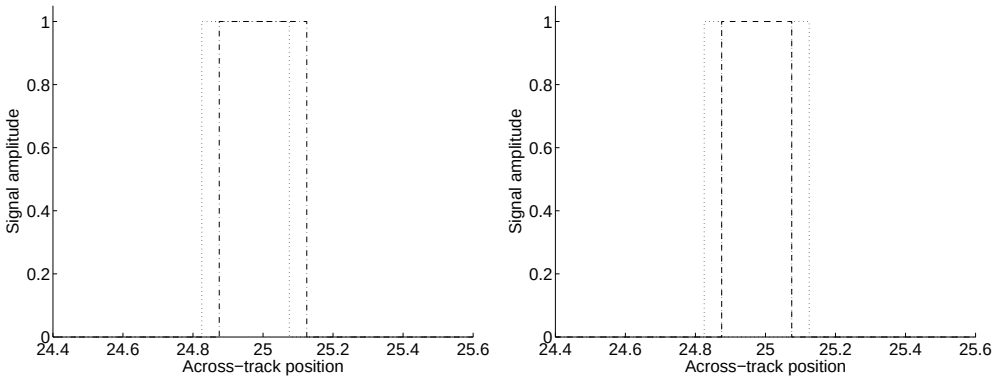
For a non-rectangular across-track resolution cell, the derivation of the equivalent $\text{SNR}_{\,m}$ of (3.16) is more difficult. The signal common to the two offset footprints must be calculated, and also the signal outside this common portion calculated. An example for both rectangular and sinc resolution cells is shown in Figure 3.4 for the partial overlap case. One point of interest is the relative sizes of the footprints of each receiver. For the lower receivers, the footprint is slightly larger than for the other receivers. However, the size of this difference is in the order of 0.1% for typical SAS systems, so can be ignored.

For the KiwiSAS parameters, the footprint shift and resulting coherence component are shown in Figure 3.5. The strong coherence loss due to footprint shift is evident in Figure 3.5, where even at an across-track range of 100 m the coherence component is 0.62 for the closely separated receivers, 0.39 for the outer pair of receivers.

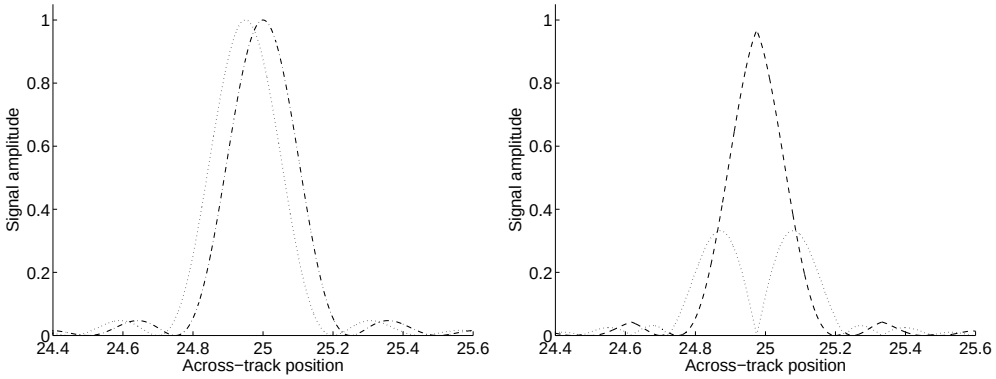
3.5 BASELINE DECORRELATION

In an interferometric sonar system, the seafloor is normally illuminated with one transmitter, and the reflected echo measured with two or more receivers with vertical separation. The positional difference of the receivers implies each receiver intersects a different part of the reflected field from the common incident wavefront. The signals received by each receiver will differ, with the expected interferometric phase difference between the receivers corrupted by the field decorrelation. The field decorrelation between the receivers is caused by a different mechanism to that of the footprint shift decorrelation, and will still be present even after perfect interpolation is used to remove the footprint shift. It is assumed here the footprint shift correction has been performed perfectly, the process of which is covered in more detail in Section 3.4. Making this co-located footprint assumption, the baseline coherence component can be calculated,

$$\gamma_b = \frac{\langle d_1^* d_2 \rangle}{\sqrt{|d_1|^2 \, |d_2|^2}}, \tag{3.19}$$

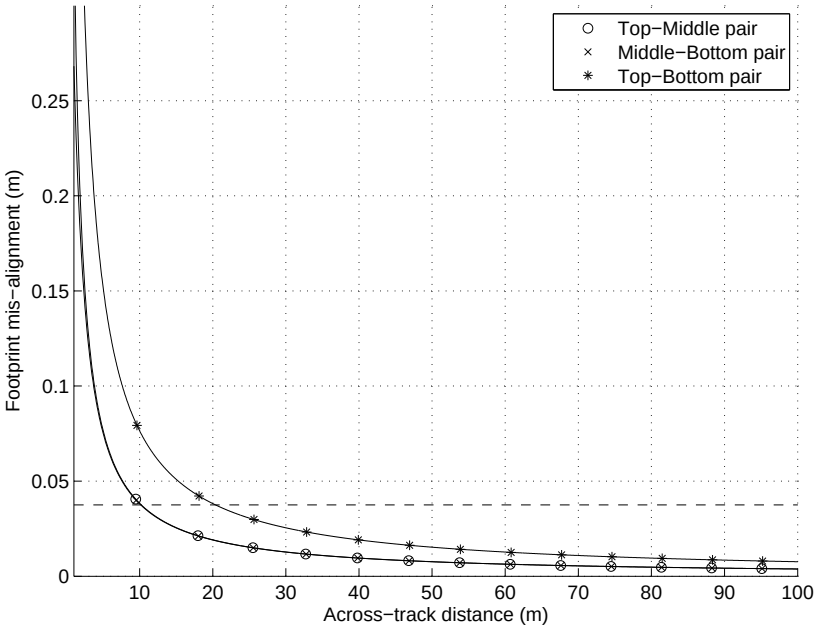


(a) Rectangular resolution cell

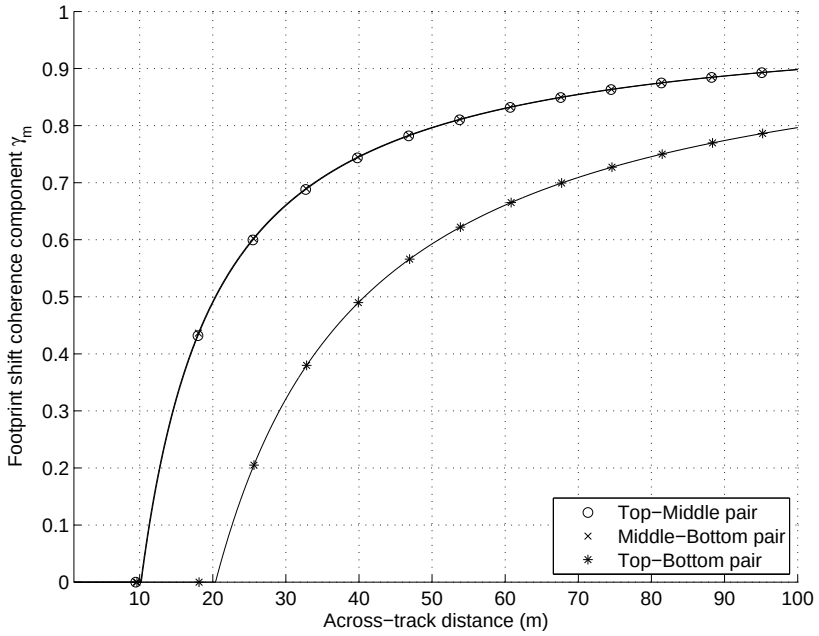


(b) Sinc resolution cell

Figure 3.4 Footprint shifts for (a) rectangular, and (b) sinc resolution cell shapes. Left plots show resolution cells for a slant-plane range of 25 m for the top receiver, and 24.95 m for the lower receiver. Right plots show the common signal between the two resolution cells (dashed), and difference (dashed). The ratio of the areas below the curves of signal and difference gives an equivalent SNR $_m$. For this example, the equivalent SNR $_m$ is 2.0 for the rectangular case, and 1.65 for the sinc case.



(a) Footprint mis-alignment



(b) Coherence components from footprint mis-alignment

Figure 3.5 Footprint shifts of the KiwiSAS system using the three interferometric pairs of receivers. Seafloor located 10 m below the sonar system, with a system resolution of 0.0375 m. The outer pair of hydrophones exhibits approximately twice the footprint misalignment of the other two hydrophone pairs because of the doubling in receiver spacing. The horizontal dashed line is the theoretical across-track resolution cell size.

where d_1 and d_2 are the signals at the two receivers. By assuming scatterers are distributed continuously and evenly over the surface, the field at a point of interest, d, can be expressed as an infinite integral,

$$d = \int_{-\infty}^{\infty} w\left(x - x_0\right) a\left(x\right) \exp\left(jk\left(r_i\left(x\right) + r_e\left(x\right)\right)\right) dx, \tag{3.20}$$

where $w\left(x - x_0\right)$ is the across-track slant-plane resolution cell shape of the imaging system, centred at an across-track position of x_0, $a\left(x\right)$ is a range dependent amplitude function, $r_i\left(x\right)$ is the range from the transmitter to the surface, $r_e\left(x\right)$ is the range from the surface to the receiver, and $k = \frac{2\pi f_0}{c}$ is the wavenumber of the pulse-compressed signal. For an InSAS system with one transmitter and multiple receivers, the two way range to the surface can be approximated as the distance to the surface from a point midway[3] between the transmitter and each receiver, known as the phase centre. For a vertical array of receivers, the array of phase centres will be a vertical array with separation of half the original receiver spacing. The phase centre approximation is valid for system geometries where the transmitter and receiver are located close to each other compared to the distance to the surface. The distance from the phase centre to the surface is,

$$r\left(x\right) = \frac{r_i\left(x\right) + r_e\left(x\right)}{2}. \tag{3.21}$$

Using the phase centre approximation, the signals at receivers a and b are given by (3.20),

$$d_1 = \int_{-\infty}^{\infty} w\left(x - x_0\right) a\left(x\right) \exp\left(j2kr_a\left(x\right)\right) dx, \tag{3.22}$$

$$d_2 = \int_{-\infty}^{\infty} w\left(x - x_0\right) a\left(x\right) \exp\left(j2kr_b\left(x\right)\right) dx\,. \tag{3.23}$$

The signals of (3.22) and (3.23) are assumed to be pulse-compressed, hence the exponential term is only dependent on the centre frequency of the incident signal. The bandwidth of the original incident signal determines the shape of $w\left(x\right)$. For a rough surface with a large number of scatterers per resolution cell, the field is a circular, complex Gaussian random variable [Goodman, 1985]. The correlation between two points of the field is defined as

$$\langle S \rangle = \langle d_1^* d_2 \rangle, \tag{3.24}$$

where $\langle \cdot \rangle$ represents ensemble averaging. It should be noted here this definition of correlation is different to that used in other research fields where a normalised version

[3] Calculated in three dimensions.

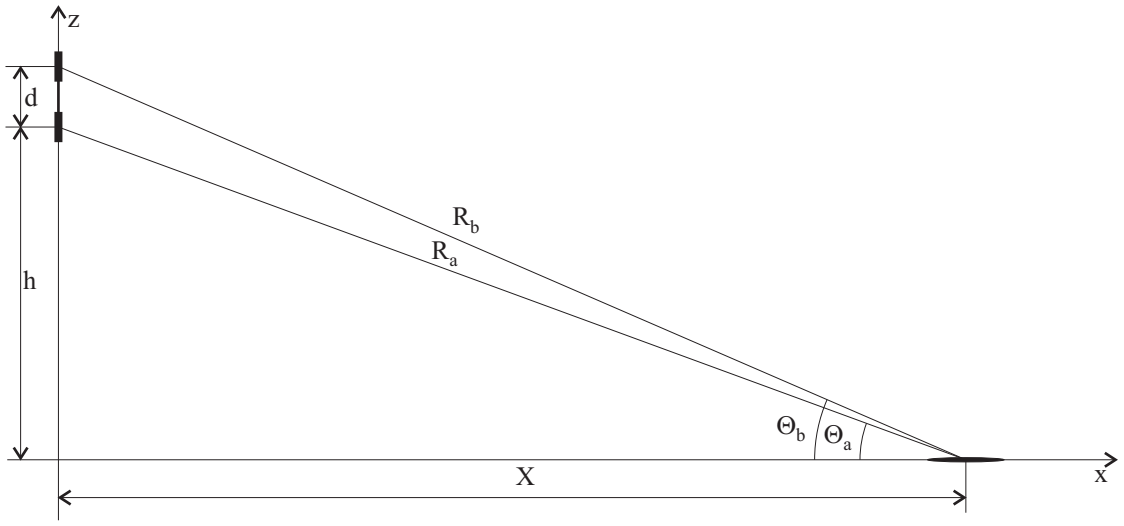


Figure 3.6 Geometry and notation of baseline decorrelation analysis. Two receivers with vertical separation d are considered in the field of a seafloor patch, centred at a across-track distance of x_0 and depth h. The radial ranges from the ground-plane to the receivers is defined as r_a and r_b, and r_{a0}, r_{b0} to the seafloor patch at across-track range x_0. Here the footprint shift effect is ignored, or assumed to be perfectly corrected.

is often used. Substituting (3.22) and (3.23) into (3.24),

$$\langle S \rangle = \left\langle \iint \begin{array}{l} w\left(x - x_0\right) a\left(x\right) \exp\left(-j2kr_a\left(x\right)\right) \\ w\left(x' - x_0\right) a\left(x'\right) \exp\left(j2kr_b\left(x'\right)\right) dx\, dx' \end{array} \right\rangle . \tag{3.25}$$

By assuming the scatterers are uniformly distributed and uncorrelated over the surface,

$$\langle a\left(x\right) a\left(x'\right) \rangle = \langle a^2 \rangle \delta\left(x - x'\right), \tag{3.26}$$

where $\langle a^2 \rangle$ is the average scatterer strength, and $\delta\left(\right)$ is the Dirac-delta function. The removal of the amplitude range dependence is valid for small changes in range, such as the differences in path lengths to each of the receivers of an InSAS system. Because of the sifting property of the Dirac delta function, the double integral of (3.25) can be reduced to a single integral over x,

$$\langle S \rangle = \langle a^2 \rangle \left\langle \int w^2\left(x - x_0\right) \exp\left(j2k\left(r_b\left(x\right) - r_a\left(x\right)\right)\right) \mathrm{d}x \right\rangle . \tag{3.27}$$

Considering the geometry and notation as shown in Figure 3.6 the distances r_a and r_b can both be expanded, for r_a giving

$$r_a = \sqrt{h^2 + x^2}. \tag{3.28}$$

At the central point of the resolution cell $x = x_0$,

$$h^2 = r_{a0}^2 - x_0^2. \tag{3.29}$$

Substitution of (3.29) into (3.28) gives

$$r_a = \sqrt{r_{a0}^2 - x_0^2 + x^2}. \tag{3.30}$$

By observing

$$(x - x_0)^2 = x^2 - 2xx_0 + x_0^2, \tag{3.31}$$

(3.30) can be written,

$$\begin{aligned} r_a &= \sqrt{r_{a0}^2 + (x - x_0)^2 - 2x_0^2 + 2xx_0} && (3.32)\\ &= r_{a0}\sqrt{1 + \frac{(x - x_0)^2}{r_{a0}^2} + \frac{2xx_0}{r_{a0}^2} - \frac{2x_0^2}{r_{a0}^2}}. && (3.33) \end{aligned}$$

Using a first order Binomial approximation[4],

$$\begin{aligned} r_a &\approx r_{a0}\left(1 + \frac{(x - x_0)^2}{2r_{a0}^2} + \frac{x_0 x}{r_{a0}^2} - \frac{x_0^2}{r_{a0}^2}\right) && (3.34)\\ &\approx r_{a0} + \frac{(x - x_0)^2}{2r_{a0}} + \frac{x_0 x}{r_{a0}} - \frac{x_0^2}{r_{a0}}. && (3.35) \end{aligned}$$

If the scattering region is small, $|x - x_0| \ll r_{a0}$, the second term of (3.35) can be neglected (Fraunhofer approximation),

$$r_a \approx r_{a0} + \frac{x_0 x}{r_{a0}} - \frac{x_0^2}{r_{a0}}. \tag{3.36}$$

Applying the cosine ratio $\cos\theta_{a0} = \frac{x_0}{r_{a0}}$,

$$\begin{aligned} r_a &\approx r_{a0} + x\cos\theta_{a0} - x_0\cos\theta_{a0} && (3.37)\\ &\approx r_{a0} + (x - x_0)\cos\theta_{a0}. && (3.38) \end{aligned}$$

Similarly, r_b can be derived, giving the pair or equations,

$$\begin{aligned} r_a &\approx r_{a0} + (x - x_0)\cos\theta_{a0}, && (3.39)\\ r_b &\approx r_{b0} + (x - x_0)\cos\theta_{b0}. && (3.40) \end{aligned}$$

[4] $(1 + x)^m \approx 1 + mx$ for $x \ll 1$.

The equations of (3.40) and (3.40) can now be substituted into (3.27),

$$\begin{aligned}\langle S\rangle &= \left\langle a^2\right\rangle \int w^2\left(x-x_0\right)\exp\left(j2k\left(r_b\left(x\right)-r_a\left(x\right)\right)\right)dx &(3.41)\\ &= \left\langle a^2\right\rangle \int w^2\left(x-x_0\right)\\ &\quad\times\exp\left(j2k\left(r_{b0}-r_{a0}+\left(x-x_0\right)\left(\cos\theta_{b0}-\cos\theta_{a0}\right)\right)\right)dx\,. &(3.42)\end{aligned}$$

Considering an offset coordinate system centred on x_0 such that

$$\begin{aligned}x' &= x-x_0, &(3.43)\\ \mathrm{d}x' &= \mathrm{d}x, &(3.44)\end{aligned}$$

(3.42) becomes

$$\begin{aligned}\langle S\rangle &= \left\langle a^2\right\rangle \int w^2\left(x'\right)\exp\left(j2k\left(r_{b0}-r_{a0}+x'\left(\cos\theta_{b0}-\cos\theta_{a0}\right)\right)\right)dx' &(3.45)\\ &= \left\langle a^2\right\rangle \exp\left(j2k\left(r_{b0}-r_{a0}\right)\right)\\ &\quad\times\int w^2\left(x'\right)\exp\left(j2kx'\left(\cos\theta_{b0}-\cos\theta_{a0}\right)\right)dx'\,. &(3.46)\end{aligned}$$

The exponential term outside the integral is the expected interferometric phase difference between the receivers, due to the difference in path length to the centre of the resolution cell $(r_{B0}-r_{A0})$. The integral term can be seen to be a Fourier integral, and can be replaced with a Fourier transform,

$$\langle S\rangle = \left\langle a^2\right\rangle \exp\left(j2k\left(r_{b0}-r_{a0}\right)\right)\mathcal{F}\left\{w^2\left(x\right)\right\}\big|_{f_x=\frac{2f_0}{c}\left(\cos\theta_{b0}-\cos\theta_{a0}\right)}\,. \tag{3.47}$$

Substitution into (3.24) and (3.19) gives an expression of the baseline decorrelation coherence component. After cancellation of the normalisation terms, the complex coherence component is,

$$\gamma_b = \exp\left(j2k\left(r_{b0}-r_{a0}\right)\right)\mathcal{F}\left\{w^2\left(x\right)\right\}\big|_{f_x=\frac{2f_0}{c}\left(\cos\theta_{b0}-\cos\theta_{a0}\right)}\,. \tag{3.48}$$

Since the interferometric phase difference is unknown, normally the degree of coherence is calculated, obtained by taking the magnitude of (3.48),

$$d_b = \left|\mathcal{F}\left\{w^2\left(x\right)\right\}\big|_{f_x=\frac{2f_0}{c}\left(\cos\theta_{b0}-\cos\theta_{a0}\right)}\right|. \tag{3.49}$$

The result is an analytical form for the baseline coherence component, dependent on the resolution cell shape $w\left(\cdot\right)$. Two commonly used resolution cell shapes, rectangular and sinc, are analysed further in the next two subsections.

3.5.1 Rectangular resolution cell

The shape of the field correlation function of (3.47) is dependent on the shape of the across-track resolution cell of the imaging system. For a narrowband, time-gated sonar system with resolution of Δx, the resolution cell is that of a rectangular function,

$$w_{\text{rect}}(x - x_0) = \begin{cases} 1 & |x - x_0| < \frac{\Delta x}{2} \\ 0 & \text{otherwise} \end{cases} . \tag{3.50}$$

Substitution of this resolution cell into the field equation of (3.47) gives the field correlation as presented in [Jin and Tang, 1996]. From (3.49) it can be seen the coherence function is the Fourier transform of the resolution shape squared. For the rectangular resolution cell used here, the baseline decorrelation function takes the form of an absolute sinc function. Of interest here is the infinite extent of the baseline decorrelation function, implying the coherence between the two interferometric receivers will be non-zero for almost any baseline separation.

3.5.2 Sinc resolution cell

A commonly signal used in SAS and InSAS systems is a linear FM chirp [Hawkins, 1996]. Upon reception the echo is matched-filtered (pulse-compressed) with the transmitted signal to reduce the signal time extent. For this matched-filter system, the expected ideal resolution cell is the autocorrelation of the transmitted signal. From (3.49) the sinc resolution cell used here, the baseline coherence component is a triangular function.

3.5.3 Critical baseline

In interferometric radar literature, the concept of a 'critical baseline' is defined. Namely, the baseline length at which the coherence between the receivers of the array is zero. Thus, the critical baseline gives a maximum possible baseline for interferometric operation, beyond which zero coherence is expected, and hence no interferometric estimates can be made. For a matched-filtered system with a nominally sinc resolution cell, this definition is unambiguous due to the triangular shape of the associated coherence function. For non-sinc resolution cell shapes, this definition of the critical baseline as defined above cannot be applied, due to the infinite extent of the coherence function. The infinite extent of the coherence function has important consequences for interferometric processing, allowing for a large separation between the receiver while still maintaining some level of coherence, albeit much reduced. It is, however, possible to define a similar concept for such resolution shapes, defined as the distance to the first zero crossing of the coherence function. This definition is less 'critical', since beyond this distance there is still significant coherence between the receivers. However, the level of coherence in

these side-lobes of the coherence function is often too low for practical phase estimation without a very large number of independent looks. For more analysis of independent looks see Section 5.3.1.

Typically, InSAS systems use a linear frequency modulated FM chirp, giving a sinc shaped resolution cell after pulse compression. The critical baseline can be approximated as[Bamler and Hartl, 1998, Rosen et al., 2000],

$$B_{\text{crit}} \approx \frac{\lambda R_s \tan\theta}{2\delta_x}, \tag{3.51}$$

where λ is the wavelength, R_s is the slant-plane imaging range, θ is the imaging angle from nadir, and δ_x is the imaging resolution in the across-track direction. As a comparison between InSAR and InSAS systems, the parameters of the ERS InSAR system and the KiwiSAS InSAS system are used, listed in Table 2.2. For the ERS parameters the critical baseline is

$$\begin{aligned} B_{\text{crit,ERS}} &\approx \frac{0.0567 \times 780 \times 10^3 \times \tan 0.40}{2 \times 9} \\ &\approx 1142\,\text{m}, \end{aligned} \tag{3.52}$$

consistent with [Bamler and Hartl, 1998]. For the KiwiSAS InSAS parameters, the critical baseline for the upper operating frequency band (100 kHz) for a point on the seafloor 50 m across-track and 10 m below the towfish is

$$\begin{aligned} B_{\text{crit,Kiwi100k}} &\approx \frac{0.015 \times \sqrt{10^2 + 50^2} \times \tan 1.37}{2 \times 0.0375} \\ &\approx 51\,\text{m}. \end{aligned} \tag{3.53}$$

For the lower frequency band of operation for the KiwiSAS (100 kHz), the critical baseline is extended to

$$\begin{aligned} B_{\text{crit,Kiwi30k}} &\approx \frac{0.05 \times \sqrt{10^2 + 50^2} \times \tan 1.37}{2 \times 0.0375} \\ &\approx 169\,\text{m}. \end{aligned} \tag{3.54}$$

The important difference between the InSAR and InSAS critical baselines can be seen when the typical operating baseline is considered as a ratio of the critical baseline,

$$B_{\text{r}} = \frac{B}{B_{\text{crit}}}. \tag{3.55}$$

The baseline ratio can be used to estimate the coherence component loss due to baseline decorrelation. For a sinc shaped resolution cell giving a triangular field correlation function (see Section 3.5.2), the coherence component is

$$\gamma_b = 1 - B_{\text{r}}. \tag{3.56}$$

For the ERS system, a typical interferometric baseline length is approximately 133 m[Gatelli et al., 1994], giving a baseline coherence component of

$$\begin{aligned} \gamma_{b,\mathrm{ERS}} &= 1 - \frac{133}{1142} \\ &= 0.8835. \end{aligned} \tag{3.57}$$

For the KiwiSAS upper frequency band, the baseline coherence component is

$$\begin{aligned} \gamma_{b,\mathrm{Kiwi100k}} &= 1 - \frac{0.155}{51} \\ &= 0.9970. \end{aligned} \tag{3.58}$$

As can be seen, the effect of baseline decorrelation for an InSAS system with the KiwiSAS parameters is negligible.

3.5.4 Transmitted signal envelope modification

In most image processing systems it is desirable to produce images with rectangular resolution cells. Thus, each resolution cell is not influenced by the neighbouring resolution cells, eliminating visually distracting side-lobes in the final image. Such a rectangular resolution cell is, however, difficult to produce in SAS systems requiring an infinite bandwidth signal, weighted with a sinc shape.

For a typical SAS system, the imaging signal used is a linear FM chirp, with matched filtering applied to the received echo. With a transmitted signal of $s(t)$ the echo signal received is

$$e(t) = as(t-\tau) + n'(t), \tag{3.59}$$

where a is an amplitude weighting, τ is the delay, and $n'(t)$ is additive noise. This has an echo spectrum

$$E(f) = aS(f)\exp(-j2\pi f\tau) + N'(f). \tag{3.60}$$

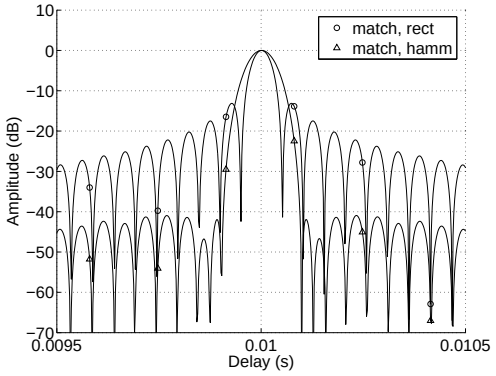
After pulse compression (matched filtering) the spectrum is

$$P(f) = E(f)S^*(f) \tag{3.61}$$

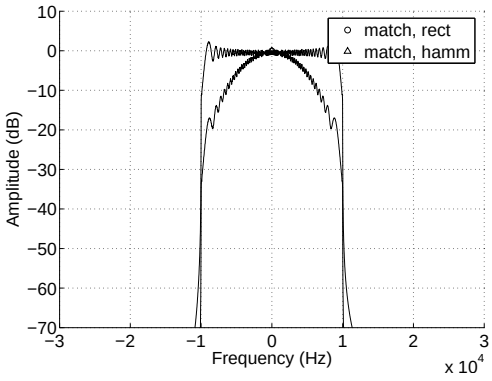
$$= a|S(f)|^2\exp(-j2\pi f\tau) + N(f). \tag{3.62}$$

Often it is desirable to modify the shape of the spectrum, primarily to decrease the energy in the side-lobes of the matched-filtering signal. Without signal filtering these side-lobes can contain significant energy, causing significants image artifacts. A first approach is to apply a frequency domain spectral window, aiming to reduce the energy contained in the time-domain sidelobes. Applying a windowing function, $W(f)$, (3.62) becomes

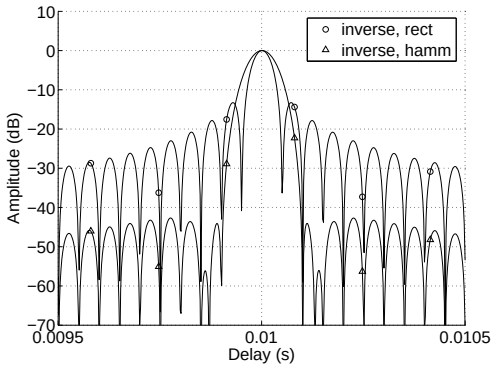
$$P(f) = E(f)S^*(f)W(f) + N(f)W(f). \tag{3.63}$$



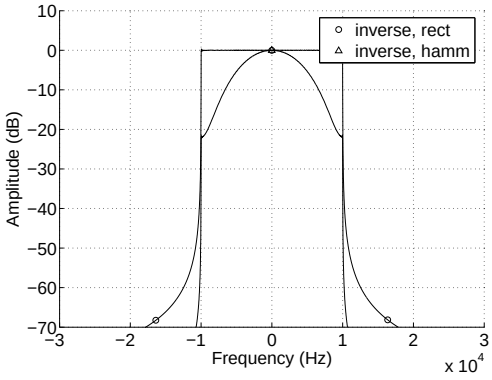
(a) Matched filtered normalised time series.



(b) Matched filtered normalised spectrum.



(c) Inverse filtered normalised time series.



(d) Inverse filtered normalised spectrum.

Figure 3.7 Modified echo time series and spectrum using: (a-b) Matched filtering, (c-d) Inverse filtering. The original echo is a 20 kHz complex chirp with a centre frequency of 0 Hz, time delayed by 0.01 s. Each plot has two different windowing functions (rectangular and hamming) applied with a window width equal to the original signal bandwidth (20 kHz). Inverse filtering decreases the effect of Gibb phenomenon as seen in the matched filter case. The application of the hamming window function can be seen to increase the width of the main lobe in the time series, decreasing resolution.

For example, the effect of applying a Hamming window can be seen in Figure 3.7(a). Without any windowing, the side-lobes can be seen to be 13 dB below the main-lobe. After the application of the Hamming window, the side-lobe suppression has been increased to greater than 42 dB. With the application of alternative windowing functions, this rejection could be further increased. The increase in side-lobe rejection does come at a cost, however, reducing the effective bandwidth of the signal. This reduction of bandwidth has the effect of widening the main-lobe width, lowering the resolution. For the example shown in Figure 3.7(a), the original 3 dB width of the main-lobe is 0.0615 m. After windowing with the Hamming window, the width is increased to 0.0915 m.

Although frequency domain windowing does provide some control over the time domain signal side-lobes, the shape is still largely influenced by the shape of the spectral original pulse-compressed signal. Any distortion in the spectrum due to pulse-

compression will still be present in the final signal, as shown in Figure 3.7(b). The passband can be seen to contain significant deviations from the ideal 0 dB. After the application of a Hamming window, these distortions are still present. An alternative windowing technique is to use an inverse filter, rather than matched-filtering. With inverse filtering the processed spectrum takes the form,

$$P(f) = \frac{E(f)\,W(f)}{S(f)}. \tag{3.64}$$

Inverse filtering has the advantage of producing no distortions in the passband, as normally seen with conventional matched-filtering. Application of a spectrum shaping window is then more effective, since the resulting spectrum will more closely match the desired window shape. In the noise free case where $E(f) = S(f)$ the shape of the final spectrum is that of the windowing function $W(f)$. The difference in spectral content between matched filtering and inverse filtering can be seen by comparing Figures 3.7(b) and 3.7(d). The matched filtered spectrum contains significant distortions in the echo passband of ± 10 kHz. In the inverse filter case, these distortions are removed, with the entire echo passband of 0 dB. Inverse filtering must be used with care; the division process of (3.64) cannot be performed if sample values of zero are present in the signal spectrum, $S(f)$. Also, if the division is performed first, energy in the resulting spectrum is no longer limited to the original bandwidth, instead covering the entire spectrum. This spectrum is then reduced when the windowing function $W(f)$ is applied, however significant aliasing effects can still occur if computed with the division within (3.64) first. If no windowing is desired, a rectangular window must still be applied, with a passband equal to the echo bandwidth. This can be seen in Figure 3.7(d), where the spectrum outside the original echo passband is attenuated. To avoid the division be zero and aliasing issues, the inverse filter can be re-formulated as

$$P(f) = E(f)\,R^*(f), \tag{3.65}$$

where $R(f)$ is the reference signal, defined as

$$R(f) = \frac{W^*(f)}{S^*(f)}. \tag{3.66}$$

When applying the windowing function, it is important to note the maximum width of the window can never exceed that of the bandwidth of the echo signal $E(f)$. If the windowing function is of greater width than the echo bandwidth, the effect is to add significant noise to the system. The difference between matched filtering and inverse filtering is less obvious when comparing the time signals of each, as shown in Figures 3.7(a) and 3.7(c). Comparing the side-lobe attenuation, inverse filtering has only a few dB improvement over matched-filtering and can reduce the SNR.

Spectral window selection

Any application of a spectral window will result in a loss in effective signal bandwidth. The bandwidth of the window must be less than the original signal and cannot introduce frequencies outside the original signal band. One commonly used windowing function is a Hamming window. This is known as a general purpose window, however, is not suited to applications where Fourier techniques are used because of its poor side-lobe rejection of -32 dB. A more suitable window for applications involving Fourier techniques is a Blackman-Harris window, because of its superior side-lobe rejection. For a 4-term Blackman-Harris window, the rejection ratio of the side-lobes is in the order of -96 dB, often in excess of the underlying SNR of the system. Thus, the energy in the side-lobes generated by the Fourier transform is below that of the original noise floor of the signal. The disadvantage of the Blackman-Harris window is the increase in the narrow baseband width, causing large resolution losses.

Display of imagery with large pixel dimensions

In a practical sonar system, one of the operational limiting factors is displaying the data to the operators. With a high resolution sonar system, the resulting image will typically contain several thousand resolution cells. When this is sampled above the Nyquist rate, the number of image pixels is often too large to display on a computer screen or printer. For example, a SAS with a bandwidth of 40 kHz and an operating swath of 100 m produces 10667 pixels, when sampled at the minimum Nyquist rate. This is far greater than any current computer monitor can display, and even beyond most printers. A computer monitor typically has a pixel-pitch of approximately 0.25 mm, so to display the full image of above a screen in excess of 2 m width is required. If the full swath image is displayed on a screen with less than the required number of pixels, the image must first be reduced in size, thus lowering the resolution. If this decimation process is done poorly[5], the resulting image often suffers badly from image artifacts. A more desirable solution is to lower the resolution and decimate the image to yield an image of the same pixels count as the display. The lowering of resolution can be achieved by lowering the bandwidth of the echo signal, as part of the inverse filtering of (3.65). The window required to achieve this bandwidth lowering is simply a narrow window with width as required, and zero outside this region. The central frequency of the windowing function can also be altered, and for sufficiently narrow windows several independent windows may be applied to the original bandwidth. Each of these multiple smaller spectral windows can then be combined in some manner, often allowing the specular nature of the data to be reduced.

[5]Often computer operating systems will implement a primitive decimation technique when down-sampling images the interest of speed, and may result in image artefacts.

3.6 TEMPORAL DECORRELATION

Temporal decorrelation causes a loss of coherence ($\gamma_t < 1$) if there is a physical change in the imaged scene during the time period between imaging of the interferometric pairs of receivers. This effect is therefore only present in a multi-pass system. For SAS, currently there are no known multi-pass systems. In the InSAR field, repeat pass interferometry is common, using different satellite orbits are used to create the interferometric baseline [Graham, 1974, Gabriel and Goldstein, 1988, Li and Goldstein, 1990, Gatelli et al., 1994, Ghiglia and Wahl, 1994, Lanari et al., 1996, Homer et al., 2002]. Also, repeat pass InSAR systems have specifically used this phenomena to map movements in the earths surface. These movements could be over a short time frame such as seasonal events [Colesanti et al., 2003], and at a much longer time frame for tectonic plate movements [Prati et al., 1990, Colesanti et al., 2003]. There has been some development on repeat pass InSAS systems [Bellec et al., 2005], but for all other conventional InSAS systems the effect of temporal decorrelation can be ignored, such that

$$\gamma_t = 1. \tag{3.67}$$

3.7 NUMERICAL PROCESSING NOISE

Almost any processing performed on the echo signals introduces processing noise to the result. A typical interferometric processing chain (see Figure 1.2) consists of many stages, each introducing noise. The typical InSAS processing chain consists of:

1. Base-banding
2. Pulse-compression
3. Motion compensation
4. Synthetic aperture reconstruction (azimuth compression)
5. Auto-focus
6. Interferometric phase estimation

The level of noise introduced by each of these steps varies depending on the algorithm, and the data format used at each stage.

3.7.1 Data format precision

The level of introduced processing noise for any algorithm has a lower limit dictated by the numerical accuracy of the data format used to represent the dataset. Whenever an operation is performed on a data value, the result cannot be stored with infinite

precision, rather an approximation is made introducing a small source of noise. If we consider this noise to be introduced by every calculation, the resulting SNR $_P$ for N_c calculations on the same value is

$$\underset{P}{\text{SNR}} \approx \frac{1}{N_c \epsilon}, \tag{3.68}$$

where ϵ is the numerical error introduced by each operation. Typically, processing platforms such as MATLAB[6] utilise double precision data formats for data storage. From the IEEE double precision data format Standard 754 [Institute of Electrical and Electronic Engineers, 1985], the limit of accuracy is given by

$$\epsilon_d = 2.2204 \times 10^{-16}. \tag{3.69}$$

Substitution into (3.68) shows the noise introduced can be ignored for values of approximately $N_c < 1 \times 10^{10}$. Note this is the number of calculations applied to a single data point, rather than the total computations applied to the whole dataset. The low value of ϵ_d allows the effect of data precision errors to be ignored for most SAS and InSAS applications.

The disadvantage of using a double precision data format is the memory requirements to store a dataset. Recently there has been increased interest in autonomous underwater vehicles (AUV) carrying SAS and InSAS systems. The physical constraints and limited electrical power resources of an AUV demand the data processing and storage sub-systems to be as small as possible, while consuming a minimum of power. Thus, it may be seen as advantageous to use a single precision data format, halving the memory requirements of a double precision format. The computation processing power is also decreased for a single precision data format, implying lower power consumption. However, caution must be used for single precision data formats, with a poorer limit of accuracy,

$$\epsilon_s = 1.1921 \times 10^{-7}. \tag{3.70}$$

Although this may initially seem a small error to be introduced, this is the maximum error introduced by every calculation made on a single data value.

3.7.2 Along-track ambiguity to signal ratio (AASR)

In a synthetic aperture sonar system, wide along-track beamwidths are produced by the small transmitter and receiver elements. The wide beam-width is reduced in extent during synthetic aperture processing. However, due to practical sonar platform operation, the synthetic aperture is sampled at some finite rate with sample spacing determined by the pulse repetition frequency (PRF), and the forward velocity of the

[6] www.mathworks.com

platform v. The finite sampling produces aliased targets in the along-track spatial frequency domain. The synthetic aperture reconstruction process incorrectly compresses these aliased targets, leading to a phenomena known as grating lobes [Hawkins, 1996]. An example of grating lobes for a single point target is shown in Figure 3.8. A complete sonar scene can be considered as a collection of many such points randomly spread over the scene with amplitudes dependent on the scene acoustic reflectivity, each with a response similar to Figure 3.8. The ratio of the point intensities to the summed grating lobe intensity is known as the along-track ambiguity to signal ratio (AASR). For traditional intensity based imagery the grating lobes are sufficiently small (low AASR) in amplitude to not cause any noticeable image degradation, however, for phase based calculations such as interferometry the grating lobe energy can have a significant effect.

Nyquist sampling theory states for an alias free signal the slowest sampling rate allowed for a band limited signal is twice the signal bandwidth. However, this strict definition of the Nyquist rate requires a band limited signal with no energy present in the spectrum beyond the signal bandwidth. Practically, this assumption is never met. Instead, a more reasonable definition is to sample the signal at twice the highest frequency. Using this higher sampling rate ensures any signal or noise present outside the desired bandwidth is not aliased back into the desired spectrum.

For synthetic aperture sonar applications, the Nyquist sampling requirement above implies the synthetic aperture must be sampled at a minimum of $D/2$ [Tomiyasu, 1978]. Although a sampling rate of $D/2$ can theoretically provide range independent resolution of $D/2$, the reconstructed image is severely degraded by the presence of along-track grating lobes of substantial energy [Callow, 2003]. These grating lobes can be considered a source of noise for interferometric applications, greatly reducing the coherence between the receiver channels. Improvements can be made by sampling the synthetic aperture at a higher rate, reducing the energy in the grating lobes, thus increasing the coherence between the interferometric receivers. The first publication regarding the along-track sample spacing requirement was by Tomiyasu in 1978 on the topic of strip-map SAR [Tomiyasu, 1978]. This paper claimed the along-track grating lobes were suppressed by the null in the along-track beam pattern for a $D/2$ sample spacing. As explained in [Hawkins, 1996], this analysis was erroneous, based on a real aperture interpretation of a synthetic aperture system. The error in analysis was due to an incorrect sequence of events, where the along-track beam-pattern effect was applied after synthetic aperture compression. [Hawkins, 1996] presents the correct interpretation where the along-track beam-pattern is applied to the raw data, before synthetic aperture compression. With the correct analysis, it becomes apparent $D/2$ sampling produces grating lobes (aliased targets) with a target strength -20 dB below the main target, as measured using the along-track ambiguity to signal ratio (AASR).

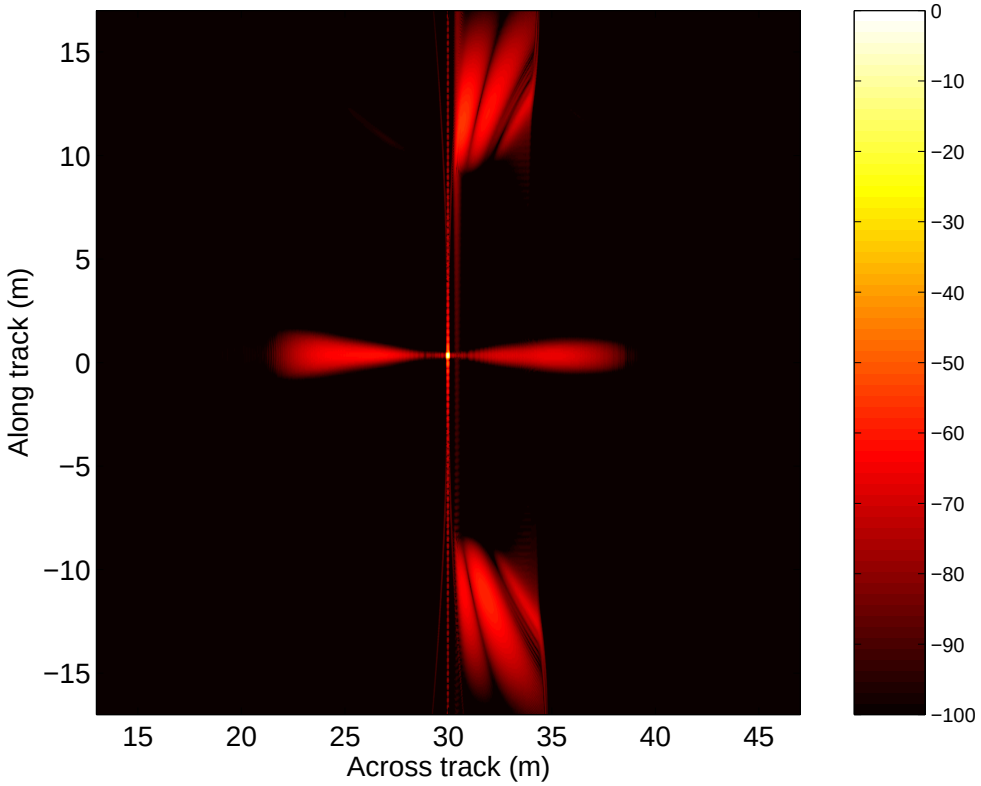
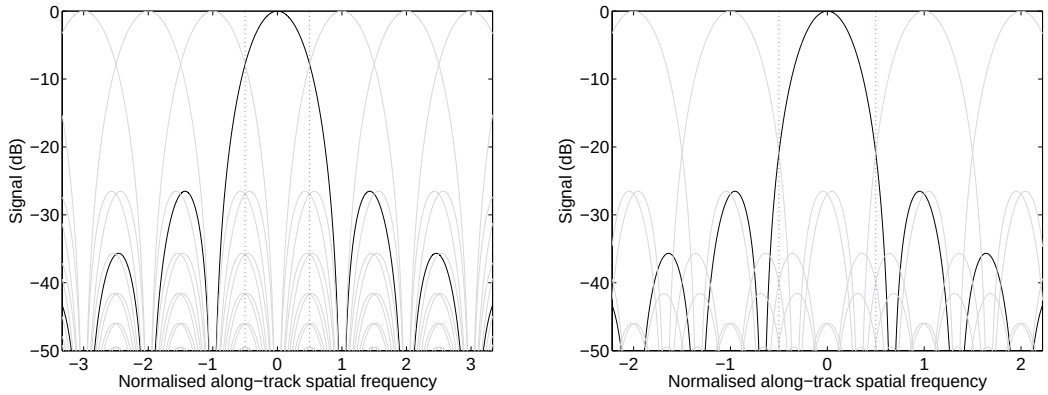
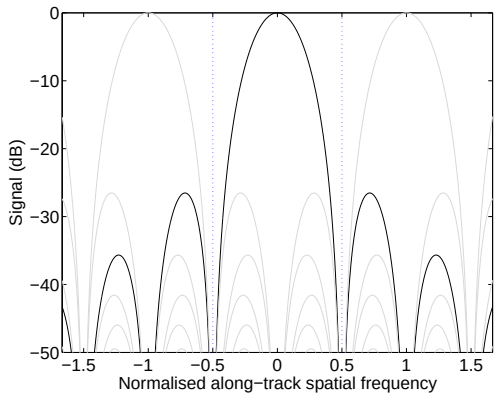


Figure 3.8 A simulated single point target at 30 m range with the KiwiSAS-IV system parameters with an along-track sample spacing of $D/3.5$, normalised to the reconstructed point target peak amplitude. Across-track pulse-compressed performed with an inverse filter using a 4-term Blackman-Harris window, reconstructed in along-track using the wavenumber algorithm. The oval shaped noise extending horizontally from the point target is the side-lobe response of the pulse-compression, the large areas of energy in along-track from the synthetic-aperture processing (grating lobes). A typical clutter sonar scene will consists of many such point responses distributed over the scene. The combined response of the scene will have a total reconstruction noise level estimated by the along-track ambiguity to signal ratio (AASR).

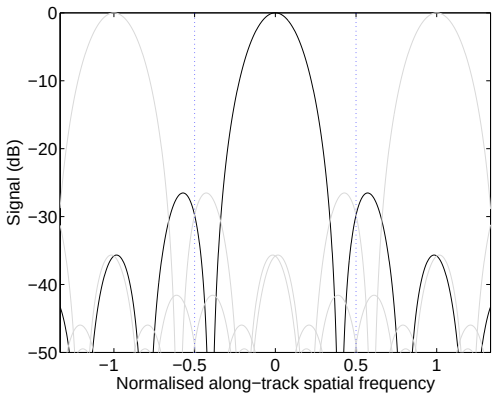


(a) $D/2$.

(b) $D/3$.



(c) $D/4$.



(d) $D/5$.

Figure 3.9 Diagrammatic explanation of the along-tack ambiguity signal to noise ratio (AASR) for an along-track sample spacing of (a) D/2, (b) D/3, (c) D/4, and (d) D/5. Transmitter and receiver elements are of equal size, D. The solid curves show the along-track spatial frequency beam-pattern, with aliased copies shown dotted. The vertical dotted lines on each plot indicate the full (100%) along-track spatial frequency bandwidth over which the synthetic aperture is formed. When the along-track sampling is poor, significant energy from the aliased copies is present within the passband (between the dotted lines) lowering the AASR.

The ratio is estimated as [Curlander and McDonough, 1991],

$$\mathrm{AASR} \approx 10 \log_{10} \left\{ \frac{\sum_{m=-\infty\, m\neq 0}^{\infty} \int_{-\infty}^{\infty} \left[W\left(\frac{k_u}{B_p}\right) A\left(k_u + m k_{us}\right) \right]^2 dk_u}{\int_{-\infty}^{\infty} \left[W\left(\frac{k_u}{B_p}\right) A\left(k_u\right) \right]^2 dk_u} \right\}, \qquad (3.71)$$

where $A\,()$ is the composite beam-pattern of the transmitter and receiver, and $W\,()$ is the along-track windowing applied. The AASR is shown graphically in Figure 3.9 for a range of along-track sampling rates.

The AASR is an important consideration for InSAS applications. Assuming the noise is independent of the signal,[7] the coherence component $\gamma_{p(a)}$ loss due to the

[7] Valid since the grating-lobe is an artefact from an scene location of significant distance from the

grating-lobes can be calculated from (3.10),

$$\gamma_{p(a)} = \frac{1}{1 + \text{AASR}^{-1}}. \tag{3.72}$$

The resulting coherence reduction for a system with equally sized, sinc beam-pattern transmitter and receiver elements is shown in Figure 3.10. At the minimum along-track sampling rate of $D/2$, the AASR is -13 dB.

The amplitude of the along-track grating lobes (giving a lower AASR) can be reduced by altering the size of the transmitter and/or receiver elements such that they are unequal in size. The transmitter to receiver size ratio (TRAR) is defined as,

$$\text{TRAR} = \frac{D_t}{D_r}. \tag{3.73}$$

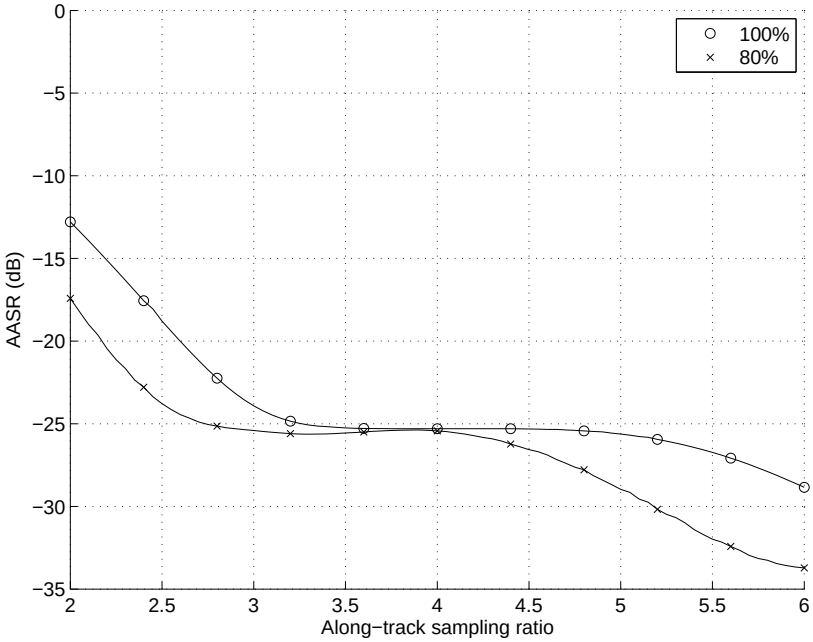
By selecting a ratio of transmitter and receiver dimensions, the beam-pattern nulls in one of the two beam-patterns can be positioned on the side-lobes of the second beam-pattern, giving an overall composite beam-pattern with much reduced side-lobes. Thus, when the signal is aliased back into the passband the noise energy is lowered, improving the AASR. The AASR for various transmitter to receiver dimension ratios is shown in Figure 3.11. For the D/2 sampling case, a size ratio of TRAR $\approx$ 1.6 yields an AASR of 0.07 (-23 dB), an improvement from -13 dB for equal sized transmitter and receiver elements (TRAR=1). The impact of this improved AASR is substantial when converted to a coherence component, improving from 0.9523 to 0.9950.

It is important to note the assumption made during this analysis of all parts of the scene having equal amplitude signal returns. While this assumption is generally true for bland, featureless sea-floors, any strongly reflecting target may violate this assumption. For example, a man-made target with defined edges will produce a strong signal with grating lobes determined by the AASR. For a sufficiently strongly reflecting target, the grating lobes may be of comparable strength to the underlying scene at the same image position as the grating lobes. In some cases the grating-lobe may even be higher in amplitude than the original scene reflectivity. In this case, the coherence of the scene will be severely reduced, giving a poor height reconstruction.

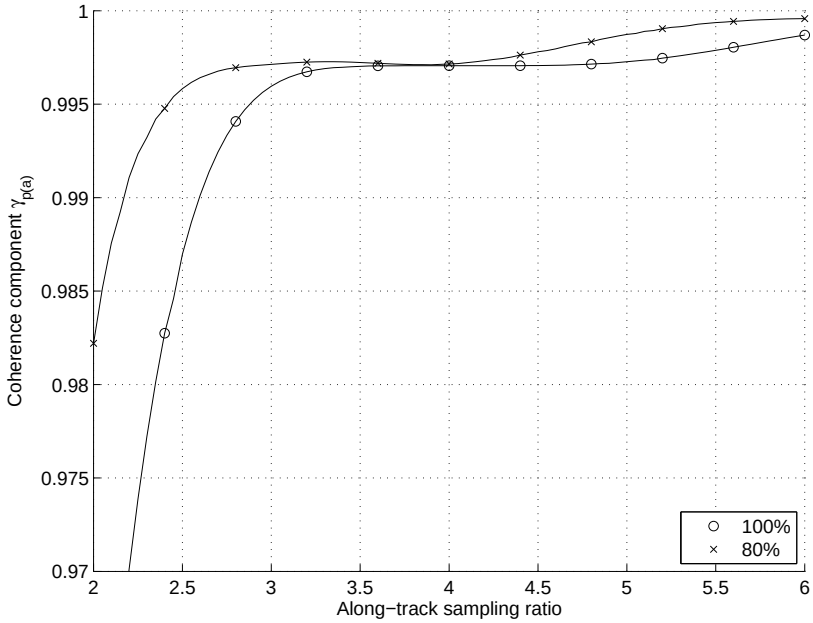
3.7.3 Interpolation noise

Almost all image processing algorithms involve some form of interpolation. The interpolation method used is, therefore, of vital importance to minimise the noise introduced to the signal. The choice of interpolation technique is critically important for any interpolation required in the phase estimation process. Any error introduced by the interpolator causes a degradation in the coherence between the channels. The

area of interest.

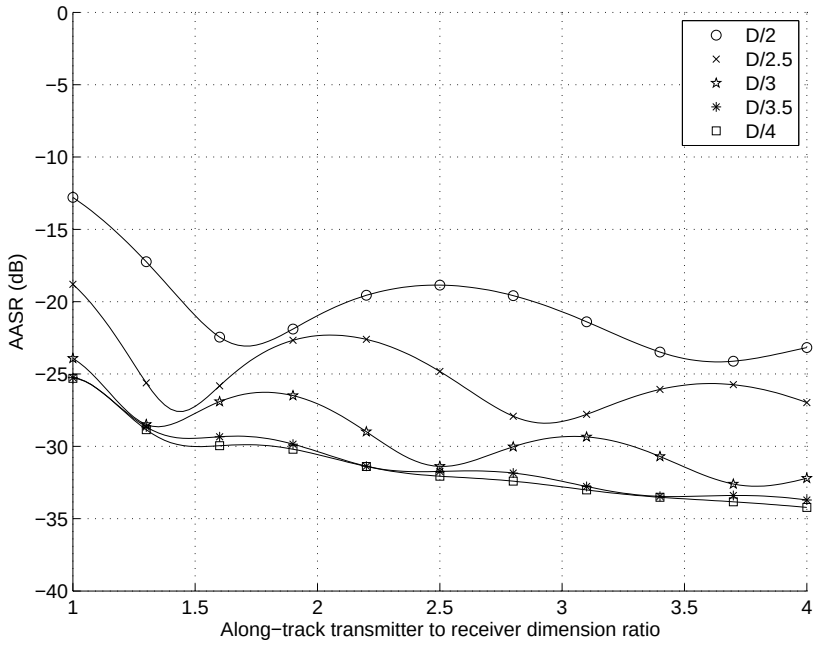


(a) AASR.

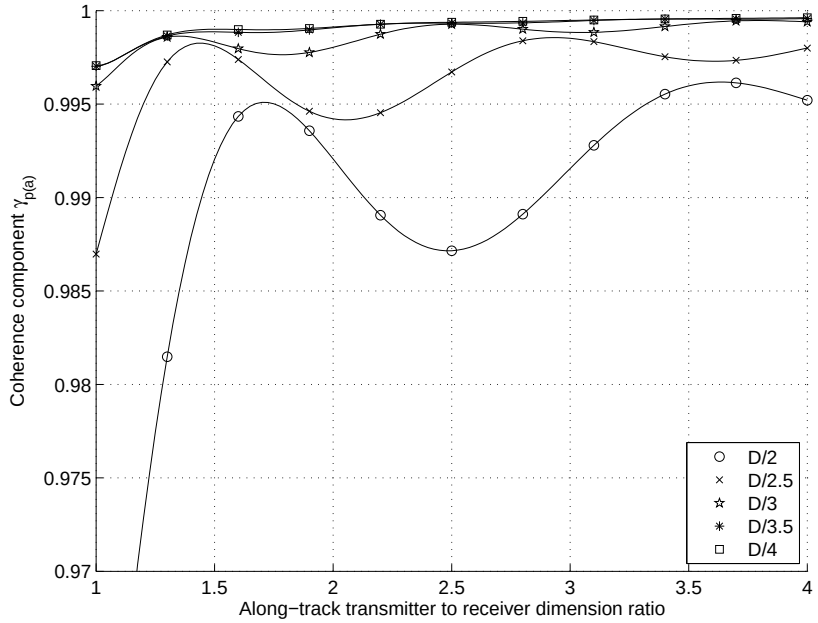


(b) Coherence component $\gamma_{p(a)}$.

Figure 3.10 The along-tack ambiguity signal to noise ratio (AASR) changes with the along-track sample spacing. Increasing the along-track sample ratio increases the Nyquist folding frequency. The AASR can be altered by weighting the aperture bandwidth, shown for a range of rectangular window widths. While weighting of the aperture bandwidth improves the AASR, the along-track image resolution is also decreased.



(a) AASR.



(b) Coherence component $\gamma_{p(a)}$.

Figure 3.11 The along-tack ambiguity signal to noise ratio (AASR) is altered as the ratio of transmitter and receiver dimensions is changed. In this figure, both transmitter and receiver have a sinc shaped beam-pattern, approximating a rectangular aperture function.

coherence loss is a function of both the interpolator kernel used, and the signal being interpolated.

An analysis of the phase noise can be undertaken by considering the spectrum of the signal and of the interpolation kernel [Parker et al., 1983, Maeland, 1988, Appledorn, 1996]. This technique was later adapted to SAR applications [Hanssen and Bamler, 1999], with interpolation kernels specifically for SAR developed based on the Knab kernel [Knab, 1983, Migliaccio and Bruno, 2003]. The analysis consists of an interpolator kernel spectrum applied to the signal to be interpolated, with replicated copies centred on multiples of the sampling frequency. The sources of noise are then: the distortion of the signal in the passband, and the aliased copies of the spectrum introduced by the non-zero spectrum of the interpolator kernel over the replicated signal spectrum. Using the system model of Figure 3.12, the interpolation has a signal to noise ratio of

$$\text{SNR} = \frac{\int_{-B/2}^{B/2} |H(f)|^2 |I(f)|^2 \mathrm{d}f}{\sum_{n\neq 0} \int_{nf_s-B/2}^{nf_s+B/2} |H(f-nf_s)|^2 |I(f)|^2 \mathrm{d}f}. \tag{3.74}$$

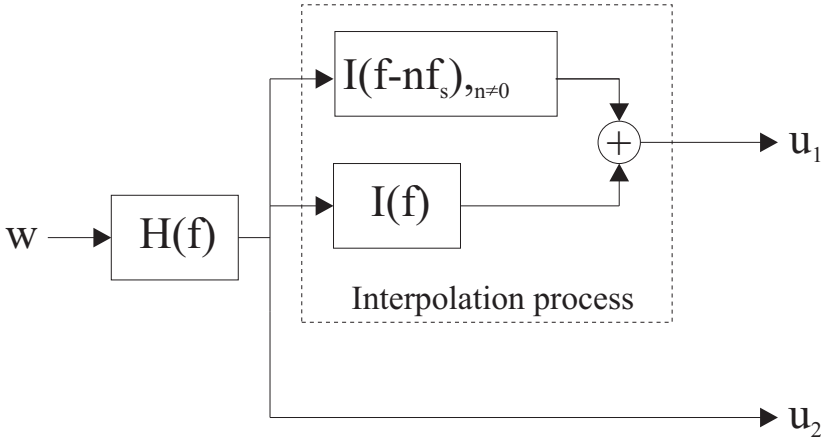
Converted to an equivalent coherence [Hanssen and Bamler, 1999],

$$\gamma = \frac{1}{\sqrt{1+\text{SNR}^{-1}}} \frac{\int |H(f)|^2 I(f) \mathrm{d}f}{\sqrt{\int |H(f)|^2 \mathrm{d}f \int |H(f)|^2 |I(f)|^2 \mathrm{d}f}}. \tag{3.75}$$

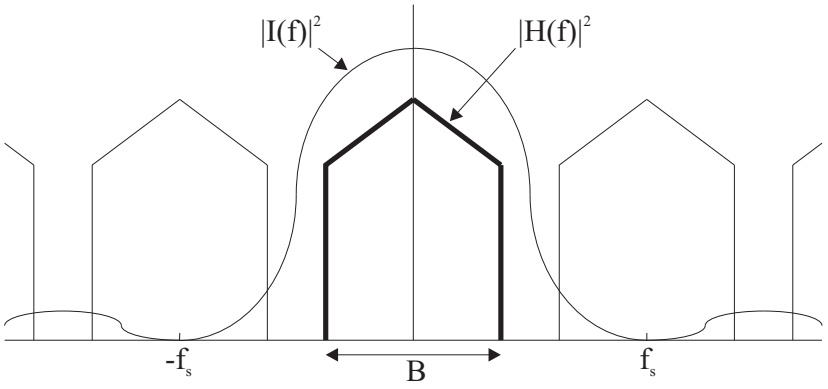
Single step interpolator kernels

A single step interpolator refers to schemes where the interpolated values are derived in one step. The output values are derived directly from the original data values, without using an intermediate step. Examples of commonly used single step interpolator kernels are: nearest neighbour, piecewise linear (often referred to as simply linear), four-point cubic, and 8-point truncated sinc. The 8-point truncated-sinc is just one example of the infinitely large set of truncated-sinc interpolator kernels. Shown in Figure 3.13 are these interpolator kernels, as well as their associated spectrum. These kernels will be used as the basis of interpolator performance, although a similar analysis can be performed on any kernel, provided the spectrum of the kernel is known.

The performance of all interpolator kernels is improved as the oversampling factor is increased. Oversampling has a two-fold effect, altering the distortion in the passband, and increasing the attenuation of the spectral aliased copies. The performance of some common interpolation kernels is shown in Figure 3.14 for a range of oversampling factors. Clearly from Figure 3.14, it can be seen the performance of a nearest neighbour interpolator is very poor, drastically decreasing coherence, even with highly oversampled data. As a result, a nearest neighbour interpolator should never be used for interferometric applications. The performance of the remaining three interpolation kernels are similar, although at low oversampling rates (< 2) the linear interpolation

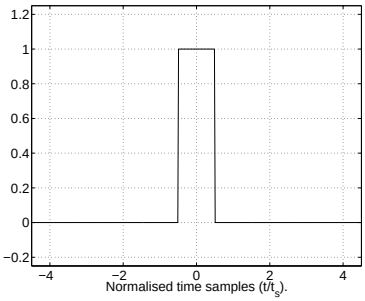


(a) System model

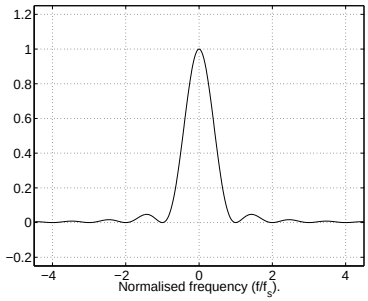


(b) Signal spectrum

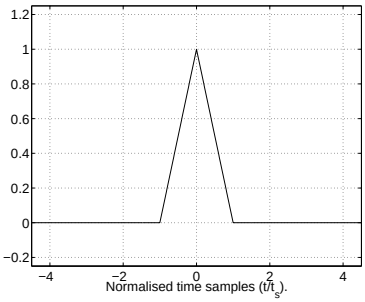
Figure 3.12 System model of signal interpolation. Input w is white noise, filtered by signal shaping function $H(f)$. The signal is then interpolated distorting the signal passband with kernel spectrum $I(f)$, and adding noise from aliased signal copies.



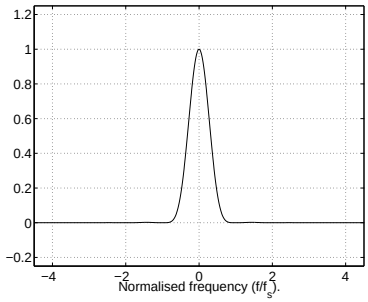
(a) Nearest neighbour kernel.



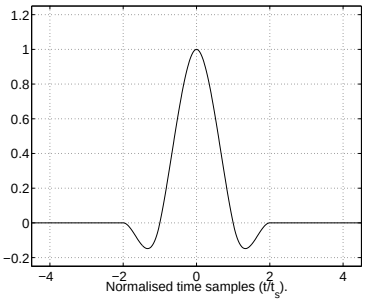
(b) Nearest neighbour kernel spectrum.



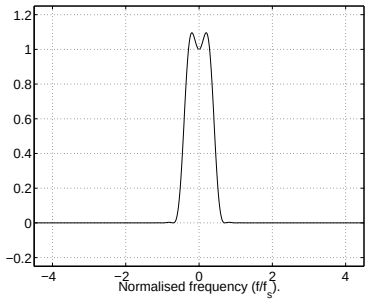
(c) Piecewise linear kernel.



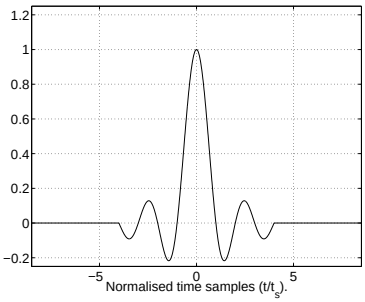
(d) Piecewise linear kernel spectrum.



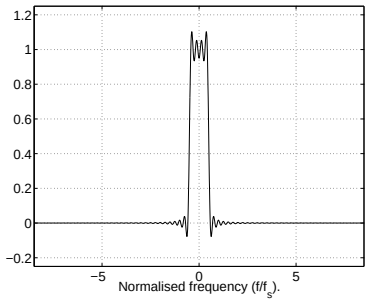
(e) Four-point cubic kernel.



(f) Four-point cubic kernel spectrum.



(g) 8-point truncated sinc kernel.



(h) 8-point truncated sinc kernel spectrum.

Figure 3.13 Examples of four interpolation kernels and their associated spectra.

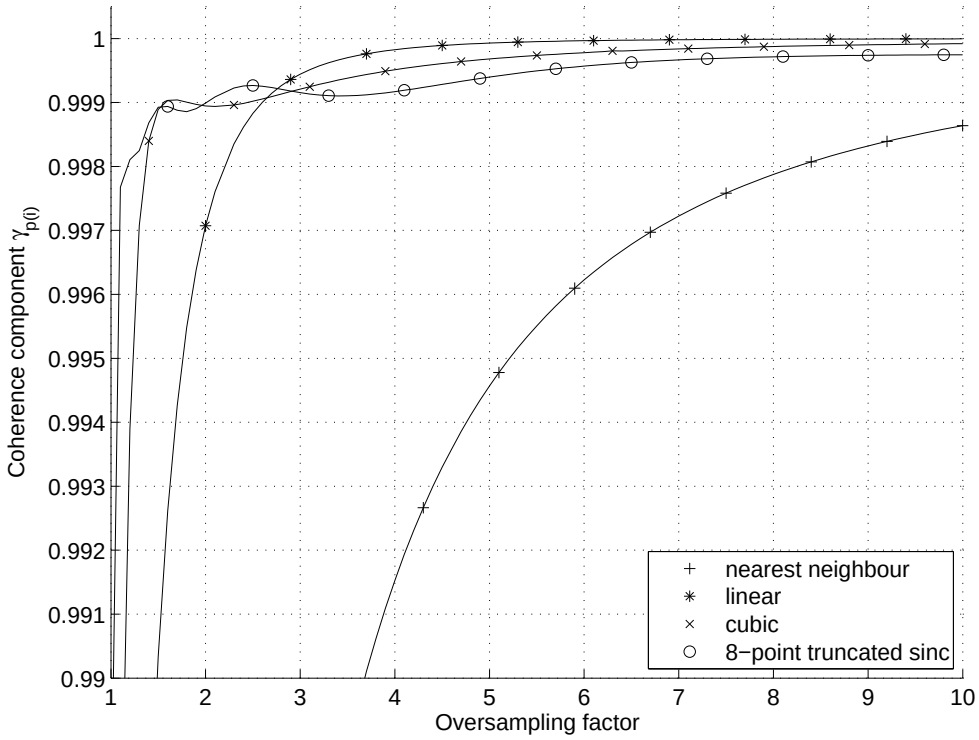


Figure 3.14 Performance (plotted as coherence) of various single stage interpolators. As the oversampling factor is increased, the coherence loss is decreased.

kernel has significantly worse performance. With datasets with an oversampling factor in excess of approximately 2.5, an interesting phenomenon occurs; the low-order linear interpolator has better performance than the higher order interpolation schemes. The cause of this performance reversal can be explained by considering the linear interpolation kernel spectrum, shown in Figure 3.13. At low oversampling rate, the rejection of aliased spectral copies of the passband are poorly rejected by the linear interpolator. Thus, much energy is aliased back into the passband, decreasing performance. Conversely, the higher order interpolators offer greater rejection of the aliased copies. At higher oversampling rates (> 2.5), the spectral aliased copies are of greater separation to the passband. In such cases, the performance is limited by the distortion introduced to the passband, rather than the rejection of the aliased spectral copies. As can be seen in Figure 3.13, the higher order interpolators have significant distortion in the passband.

For InSAS applications, the sampling rate is normally a fixed parameter, hence the advantages of oversampled data cannot be exploited. Typically, InSAS systems will sample at a rate barely above the Nyquist folding frequency; high order interpolation kernels should be used for all interpolation performed. The disadvantage of using higher order interpolation schemes is increased computational load, however, for InSAS applications the benefits of a higher data coherence warrant the extra computation

required.

Multiple step interpolator kernels

For InSAS applications, the initial sampling rate is normally only slightly oversampled, often with an oversampling factor in the region of 1.5–2.0. From Figure 3.14 it is evident such an oversampling rate requires the use of a higher order interpolation kernel. Application of a high order kernel for every desired output point is computationally expensive. Instead, fast techniques (such as Fourier based interpolation) can be used, however, these algorithms cannot be used to provide arbitrarily spaced output samples. To counter this problem, the concept of a two-stage interpolation scheme is used. Rather than interpolating the input data directly in one step, two stages of interpolation are used. Firstly, a fast, high order interpolation scheme is used to re-sample the data to a higher, albeit fixed, rate. Secondly, the up-sampled data is then interpolated to the desired sample location using a lower order interpolation kernel. The second stage has a much increased oversampling factor, allowing the use of a fast, low order interpolation kernel for the final interpolation to arbitrary sample positions. As discussed in Section 3.7.3, it is advantageous to use a low order interpolation kernel for this second stage, with lower distortion with highly oversampled data. The main advantage of a two stage scheme is the computation gains made in comparison to a single stage interpolation scheme with a high order kernel.

3.7.4 Coherence processing gains

The coherence of an InSAS system is governed by the coherence components outlined in Section 3.2. These five coherence components combine to give the overall expected coherence of the InSAS data at time of collection. By applying corrective processing algorithms, some of the coherence components can be increased, increasing the the overall coherence of the dataset. The coherence components can be categorised into two subsets; those which can be increased by corrective processing, and those which cannot be. Using this classification gives coherence components not able to be increased:

- baseline decorrelation, γ_b,
- acoustic noise, γ_n,
- temporal decorrelation, γ_t,

and components able to be increased,

- footprint mis-alignment, γ_m,
- processing noise, γ_p.

Of note here is the classification of temporal decorrelation as an uncorrectable factor. Although uncorrectable, the usual value of this coherence component is $\gamma_t = 1.0$, implying no correction is necessary anyway. For a repeat-pass system the coherence component would be lower ($\gamma_t < 1$).

Chapter 4

TIME DIFFERENCE ESTIMATION OF INTERFEROMETRIC SIGNALS

The crucial step in all interferometric systems is estimating the time delay between the incoming wavefronts across the elements of the interferometer. Considering a single interferometric pair of receivers, the echoes received can be modelled as

$$d_1(t) = e(t) + n_1(t), \tag{4.1a}$$

$$d_2(t) = \alpha e(t-\tau) + n_2(t), \tag{4.1b}$$

where $e(t)$ is the noise-free echo response, α models amplitude differences between the echoes, and $n_1(t)$, $n_2(t)$ are the additive noise components. The goal of the time delay algorithm is to estimate the time delay τ, given the two measurements d_1,d_2. For a multiple receiver setup, with multiple interferometric receiver pairs, each pair can be modelled separately using (4.1).

For a narrowband system, the time delay τ of (4.1) will have a phase-only affect on the measurements. Thus, the time delay can be estimated from the phase difference between the measurements, found simply as the angle of the Hermitian product of the two signals,

$$\phi = \angle(d_1 d_2^*), \tag{4.2}$$

where * denotes the complex conjugate operator. The time delay τ can then be inferred from the centre frequency of the system. Due to the circular nature of the phase of complex signals, phase difference estimation between two signals d_1 and d_2 can normally only be estimated modulo 2π. Recovery of the absolute phase difference can often be achieved using two dimensional phase unwrapping techniques, seeded with prior information of a reference point within the scene. In situations with more than two receivers (giving more than one interferometric estimate), multiple interferometric phase estimates can be made, and must be combined to give one overall phase difference estimate.

For a broadband system, the effect of τ cannot be considered as phase only, since a given time delay will result in different equivalent phase changes across the signal

spectrum. Hence for a broadband system, direct phase differences on a per-pixel basis cannot be used to give an estimate of the time difference τ. Instead, τ must be estimated using an extended portion (multiple pixels) of the signals around the point of interest.

4.1 GENERALISED TIME DELAY ESTIMATION

A generalised time delay estimate can be made by finding the maximum in the cross correlation of the two signals of (4.1). The time delay estimate $\widehat{\tau}$ is

$$\widehat{\tau} = \arg\max_{\tau} \left\{ \hat{R}_{d_1 d_2}(\tau) \right\}, \tag{4.3}$$

where

$$R_{d_1 d_2}(\tau) = \mathrm{E}\left\{ d_1(t)\, d_2^*(t-\tau) \right\}. \tag{4.4}$$

The expectation operator of the cross correlation can be estimated using a portion of data around the point of interest,

$$\hat{R}_{d_1 d_2}(\tau) = \begin{cases} \frac{1}{T} \int_{\tau}^{T} d_1(t)\, d_2^*(t-\tau)\, dt & \tau \geq 0 \\ \hat{R}^*_{d_1 d_2}(-\tau) & \tau < 0 \end{cases}, \tag{4.5}$$

where T is the observation length over which the correlation is performed. This estimator is biased, substitution of the divisor T with $T - |\tau|$ removes the bias. However, the biased estimator as stated is often preferable since it has a lower mean-squared error than the unbiased estimator[Jenkins and Watts, 1968]. The estimate of the time delay $\widehat{\tau}$ of (4.3) will exhibit some variation about the true time delay τ. A lower bound on the variance of the estimate error can be made using the the Cramér-Rao lower bound (CRLB) [Knapp and Carter, 1976], given by

$$\mathrm{Var}\left[\widehat{\tau}\right] \geq \left[2T \int_0^{\infty} (2\pi f)^2 \frac{|\gamma_{ab}(f)|^2}{1 - |\gamma_{ab}(f)|^2}\, df \right]^{-1}, \tag{4.6}$$

where $\gamma_{ab}(f)$ is the coherence between signals d_1 and d_2 at each frequency. The coherence has been shown to give an effective overall SNR of [Carter et al., 1973],

$$\mathrm{SNR}\ (f) = \frac{|\gamma_{ab}(f)|}{1 - |\gamma_{ab}(f)|}, \tag{4.7}$$

thus

$$\frac{|\gamma_{ab}(f)|^2}{1 - |\gamma_{ab}(f)|^2} = \frac{1}{2} \left(\frac{1}{\mathrm{SNR}\ (f)} + \frac{1}{2\,\mathrm{SNR}^{\,2}(f)} \right)^{-1}. \tag{4.8}$$

If the signal and noise components of the SNR have an equal power spectra ratio, and are band limited over bandwidth B centred at f_c,

$$\text{SNR} = \text{SNR}\ (f)\,. \tag{4.9}$$

Substitution of (4.8) and (4.9) into (4.6) yields

$$\text{Var}\left[\widehat{\tau}\right] \geq \frac{1}{(2\pi f_c)^2}\frac{1}{TB}\frac{1}{1+\frac{B^2}{12f_c^2}}\left[\frac{1}{\text{SNR}}+\frac{1}{2\,\text{SNR}^{\,2}}\right]. \tag{4.10}$$

For more detail on this result see [Fortune, 2005].

4.1.1 Complex baseband signal correlation

When dealing with signals of low bandwidth around a high centre frequency (high Q signal), it is often impractical to store and process the signal in its raw, real form. More often the signal is down-modulated to a complex baseband signal, thus reducing the Nyquist sampling requirements greatly. Four correlation strategies exist for a pair of complex baseband signals: re-modulation to the original centre frequency, magnitude only correlation, phase only correlation, and phase based fine estimation from a magnitude only coarse estimation. The first of these four methods yields a CRLB as defined by (4.10), the remaining three methods yielding higher CRLBs. The four strategies can be seen pictorially in Figure 4.1 where the curvature of the correlation around the peak can be seen to be different for each case; higher curvature giving a lower CRLB.

Magnitude-only correlation

Using just the magnitude (envelope) of the correlation gives a time delay estimate,

$$\widehat{\tau}_m = \arg\max_{\tau}\left\{\left|\hat{R}_{d_1d_2}\left(\tau\right)\right|\right\}. \tag{4.11}$$

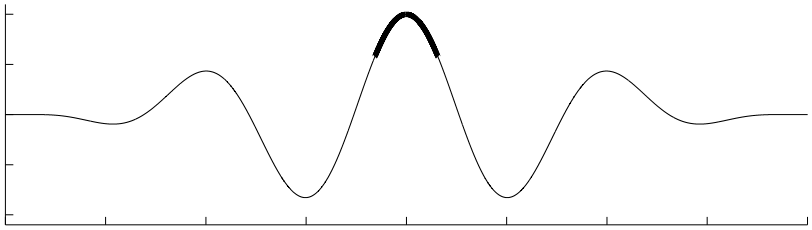
The CRLB of this estimator can be found from (4.10) by removing the centre frequency term yielding,

$$\text{Var}\left[\widehat{\tau}_m\right] \geq \frac{1}{4\pi^2}\frac{1}{TB}\frac{12}{12+B^2}\left[\frac{1}{\text{SNR}}+\frac{1}{2\,\text{SNR}^{\,2}}\right]. \tag{4.12}$$

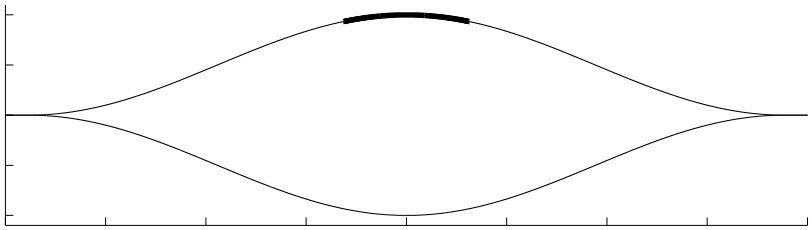
Phase-only correlation

Using just the phase of the complex baseband correlation, a time delay estimate can be formed,

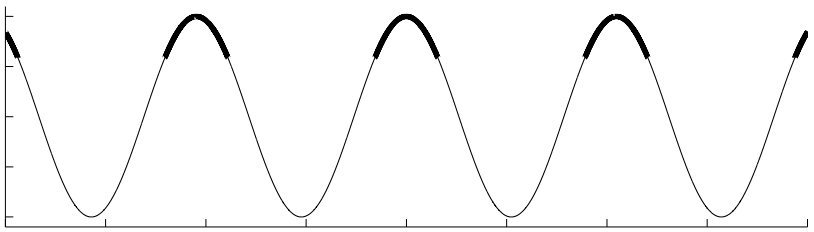
$$\widehat{\tau}_p = \frac{\arg\left(\left\{\hat{R}_{d_1d_2}\left(\tau\right)\right\}\right)}{2\pi f_c}. \tag{4.13}$$



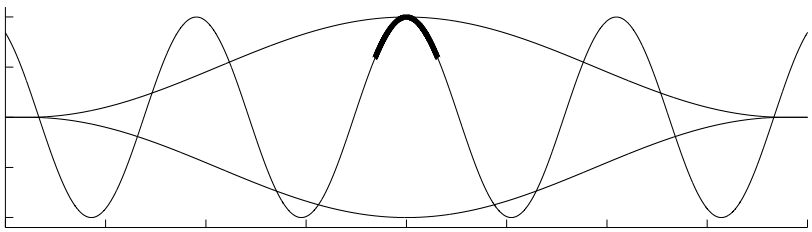
(a) Modulated real correlation



(b) Magnitude-only complex correlation



(c) Phase-only complex correlation



(d) Magnitude then phase (quasi-narrowband) complex correlation

Figure 4.1 Correlation peak estimation using (a) modulated real correlation, (b) magnitude-only complex correlation, (c) phase-only complex correlation, (d) magnitude then phase quasi-narrowband complex correlation combining (b) and (c). Bold line segments indicate curvature around estimate; higher curvature yields a lower variance estimate error. The lowest variance is exhibited by (a) but is computationally expensive. (c) exhibits similar performance to (a) but suffers from ambiguous 2π phase wraps. (d) achieves the same variance performance as (c), with the ambiguity removed using (b) as a coarse estimate.

Using a phase-only approach is equivalent to an ultra narrowband signal with zero bandwidth around the centre frequency. Hence the CRLB can be found from (4.10) by setting the bandwidth term to zero (this does not apply to the time-bandwidth term),

$$\mathrm{Var}\left[\widehat{\tau}_p\right] \geq \frac{1}{(2\pi f_c)^2}\frac{1}{TB}\left[\frac{1}{\mathrm{SNR}} + \frac{1}{2\,\mathrm{SNR}^{\,2}}\right]. \tag{4.14}$$

This lower bound is close to the original full correlation of (4.10), however, can exhibit modulo-2π phase wrapping errors for time delays equivalent to multiple wavelengths. This can be seen in Figure 4.1 where the curvature is similar to that of the full correlation, however the choice of peak can be ambiguous.

Magnitude then phase correlation (Quasi-narrowband)

The most accurate (lowest CRLB) complex correlation technique is to use the phase of the correlation, yielding a CRLB given by (4.14). However, the time-delay estimate from a phase only approach can suffer from 2π ambiguities. To overcome this shortcoming, a two step procedure can be used. First the time delay is estimated using a magnitude only approach, then refined using a phase only estimate. This approach is sometimes referred to as a quasi-narrowband approach [Shippey et al., 1998] and has been used for several algorithms within the SAS field [Shippey et al., 1998, Callow et al., 2001, Wang et al., 2001, Bellettini and Pinto, 2002]. Assuming the first (coarse) delay estimation yields an estimate with error of less than $\frac{1}{4\pi f_c}$, the final (fine) delay estimate will be unambiguous. In this case, the CRLB of the final estimate will be the same as for the phase-only approach,

$$\mathrm{Var}\left[\widehat{\tau}_q\right] \geq \frac{1}{(2\pi f_c)^2}\frac{1}{TB}\left[\frac{1}{\mathrm{SNR}} + \frac{1}{2\,\mathrm{SNR}^{\,2}}\right]. \tag{4.15}$$

If the first stage coarse time delay does not overcome the ambiguity problem of the phase only approach, the fine time delay estimate is erroneous. The performance of the quasi-narrowband approach approaches that of the full real correlation and is normally preferable because of the lower computational and memory storage requirements.

4.2 INTERFEROMETRIC PHASE IMPLICATIONS OF COHERENCE

Due to the side-scan geometry of an InSAS system, a small error in the phase difference estimation between the receivers produces a large error in the height estimation. Therefore, it is important to make an accurate estimate of the true phase difference. The coherence of the interferometric channels can be used as a measure of the overall interferometric signal quality, from which an estimate of the height error can be

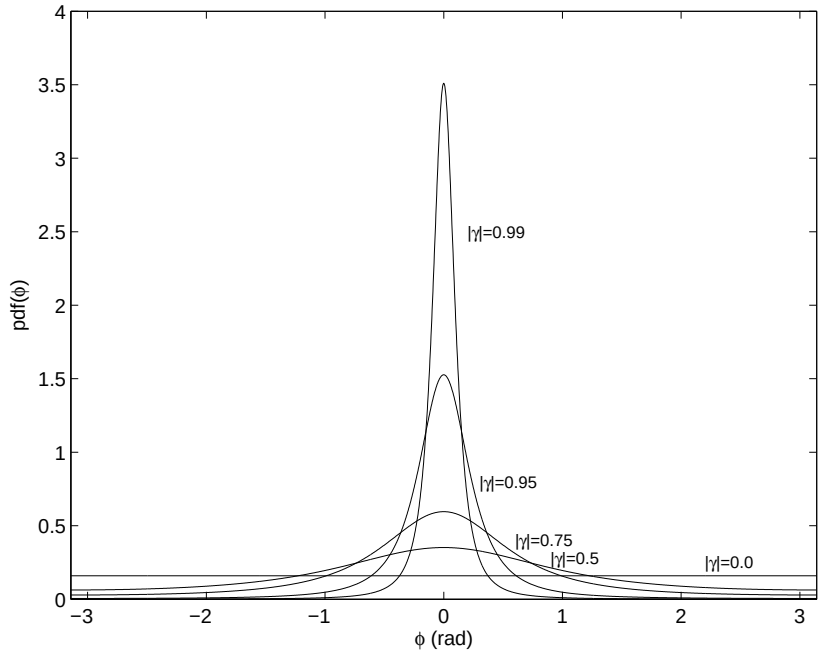


Figure 4.2 Probability distribution of phase with differing degrees of coherence for a true phase difference of $\phi_0 = 0$.

calculated. Due to the stochastic nature of SAS signals, of particular interest for interferometric calculations is the probability density function (pdf) of the phase between the two signals for a given level of coherence. For unity coherence ($\gamma = 1$), the phase difference between the interferometric pair is noise free, hence the pdf will be a Dirac delta function located at the true phase difference. At the other extreme, for complete loss of coherence ($\gamma = 0$), no phase difference information can be obtained, yielding a uniform pdf across all phase differences. The pdf of the interferometric phase has been shown to be [Davenport and Root, 1987, Just and Bamler, 1994],

$$\begin{aligned} \text{pdf}\left(\phi\right) &= \frac{1-|\gamma|^2}{2\pi}\frac{1}{1-|\gamma|^2\cos^2\left(\phi-\phi_0\right)} \\ &\quad \times\left\{1+\frac{|\gamma|\cos\left(\phi-\phi_0\right)\arccos\left[-|\gamma|\cos\left(\phi-\phi_0\right)\right]}{\left[1-|\gamma|^2\cos^2\left(\phi-\phi_0\right)\right]^{1/2}}\right\}, \end{aligned} \tag{4.16}$$

where ϕ is the estimated phase difference, and ϕ_0 is the true phase difference. The phase difference ϕ is periodic modulo-2π, so only needs to be considered over the range $\pm\pi$, as shown in Figure 4.2. An important consequence of (4.16) is the variance of the estimated phase for a given coherence; the variance of the height estimate is directly related to the variance of the phase difference estimation. The phase difference variance

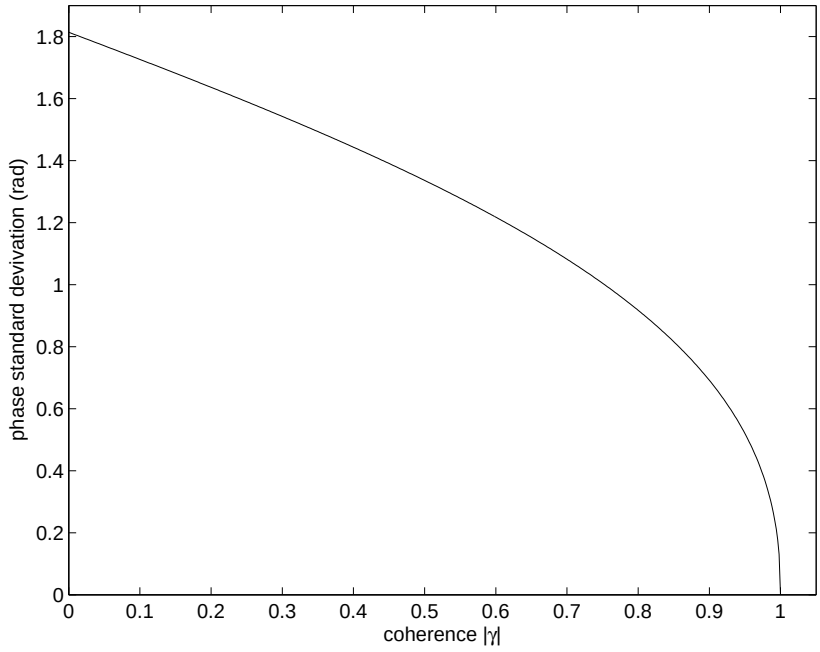


Figure 4.3 Estimated phase difference standard deviation from the true phase difference increases rapidly as the degree of coherence reduces from unity. Typical InSAS geometry magnifies any phase difference errors giving poor accuracy in the height estimate.

can be found from the second moment of the pdf [Bamler and Just, 1993],

$$\mathrm{Var}\left[\phi\right] = \int_{-\pi}^{\pi} \phi^2 \mathrm{pdf}\left(\phi + \phi_0\right) \mathrm{d}\phi. \tag{4.17}$$

As can be seen in Figure 4.3, the phase difference standard deviation rapidly increases as the coherence reduces from $\gamma = 1$. It is therefore important to maintain a high level of coherence through careful selection of system and processing parameters, and correct for any coherence loss where possible.

4.3 PHASE DIFFERENCE ESTIMATION FROM MULTIPLE LOOKS

In Section 4.2 only the one look ($L = 1$) case is considered, giving one estimate of the true phase difference. In the one look case, any slight lowering of coherence from $\gamma = 1$ causes a large increase in the variance of the phase difference estimate error, as given by (4.16) and shown in Figure 4.3 . One technique to reduce the phase difference estimate sensitivity for non-unity coherence is to combine multiple looks ($L > 1$) of the same point. The pdf for L independent looks [Lee et al., 1993, Bamler and Hartl,

1998] is

$$\begin{aligned} \text{pdf}\left(\phi, L\right) &= \frac{\Gamma\left(L+\frac{1}{2}\right)\left(1-|\gamma|^2\right)^L |\gamma| \cos\left(\phi-\phi_0\right)}{2\sqrt{\pi}\Gamma\left(L\right)\left(1-|\gamma|^2\cos^2\left(\phi-\phi_0\right)\right)^{L+\frac{1}{2}}} \\ &\quad + \frac{\left(1-|\gamma|^2\right)^L}{2\pi} {}_2F_1\left(L, 1; \frac{1}{2}; |\gamma|^2\cos^2\left(|\phi-\phi_0|\right)\right), \end{aligned} \tag{4.18}$$

where $\Gamma\left(\cdot\right)$ is the Gamma function, and ${}_2F_1\left(\cdot\right)$ is the hypergeometric function. Plotted in Figure 4.4 is the multi-look pdf for a fixed level of coherence, the pdf can be seen to be tending towards the ideal Dirac-delta shape. The phase difference variance for the multi-look case can be found using (4.17), plotted Figure 4.5 as standard deviation. Since the calculation of an analytic expression for the phase difference variance yields an unwieldy expression, often the CRLB is used as an approximation. The CRLB of the multi-look phase difference variance has been described in [Rodriguez and Martin, 1992] as

$$\text{Var}\left[\phi_L\right] = \frac{1-|\gamma|^2}{2L|\gamma|^2}. \tag{4.19}$$

This result can be derived by substitution of (3.10) into (4.15) and substitution of $L = BT$ at $\phi = 2\pi f_c \tau$. It has been shown in [Fortune, 2005] that the approximation is good for a large number of looks and high coherence. Under these conditions the probability of forming an estimate with an incorrect number of 2π phase-wraps is minimised. [Fortune, 2005] states the approximation is poor for $L < 8$ or $|\gamma| < 0.2$.

Multiple looks of the same scene can be obtained in several ways including; multiple interferometric receiver pairs, multiple passes of the scene, orthogonal imaging signals, azimuth filtering, multiple frequency bands (at time of data acquisition), and sub-banding (post processing of original bandwidth into smaller sub-bands). Orthogonal imaging signals are often used in InSAR systems using signal polarizations, however polarization is not possible for InSAS applications. Multiple frequency bands and sub-banding are similar, differing only by when the frequency sub-division occurs; either at time of data acquisition for multiple frequency bands, and during post-processing in the case of sub-banding. The effect of these two methods on the image resolution is different as discussed in Section 5.3.1. These techniques can be used individually or in combination to produce many extra looks of the scene. However, to obtain improvements in coherence as given by (4.19), all the looks must be statistically independent. In practice, this is often not the case since the looks will be partially correlated. Thus, a concept of effective looks (L_E) is used, based on the degree of correlation between the looks where,

$$L_E \leq L. \tag{4.20}$$

An example is using an array of three equally spaced vertically separated receivers.

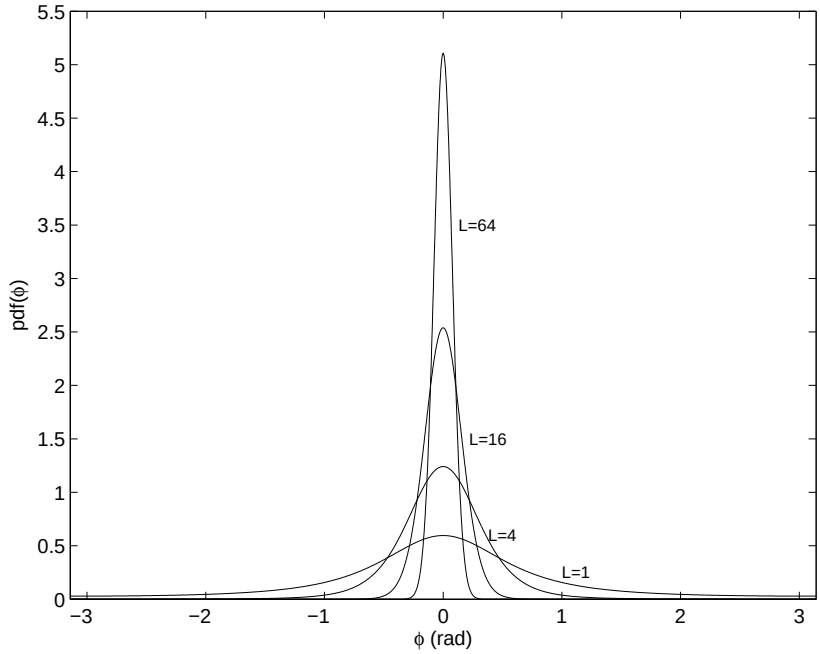


Figure 4.4 Probability distribution of phase with differing number of independent looks (L) for a true phase difference of $\phi_0 = 0$. As the number of looks is increased the pdf tends towards a delta centred on the true phase difference ϕ_0. Coherence in all cases is $|\gamma| = 0.75$.

Although this array gives three interferometric signal pairs, they are not independent ($L_E \neq 3$). Instead, as shown in Section 5.3.1, an ideal three element interferometric array will provide $L_E = 2$ looks.

4.4 MAXIMUM LIKELIHOOD PHASE DIFFERENCE ESTIMATION

The maximum likelihood (ML) phase difference from a set of data points gives the best estimation of the phase difference of the data points. For an array of N_h vertically separated receivers with an incoming planar wavefront, the received signal at receiver h can be modelled as

$$d_h = a_h m_h \exp\left(j\frac{h-1}{N_h}\phi\right) + n_h, \tag{4.21}$$

where a_h is the signal amplitude, m_h models the multiplicative noise of speckle, n_h models all additive noise, and ϕ is the phase difference as a result of the receiver separation. This model also assumes the system is narrowband, giving a phase only change between the receivers with no time difference. Assuming equal amplitude across all signals ($a = a_h$), each with a Gaussian distribution, the likelihood function is

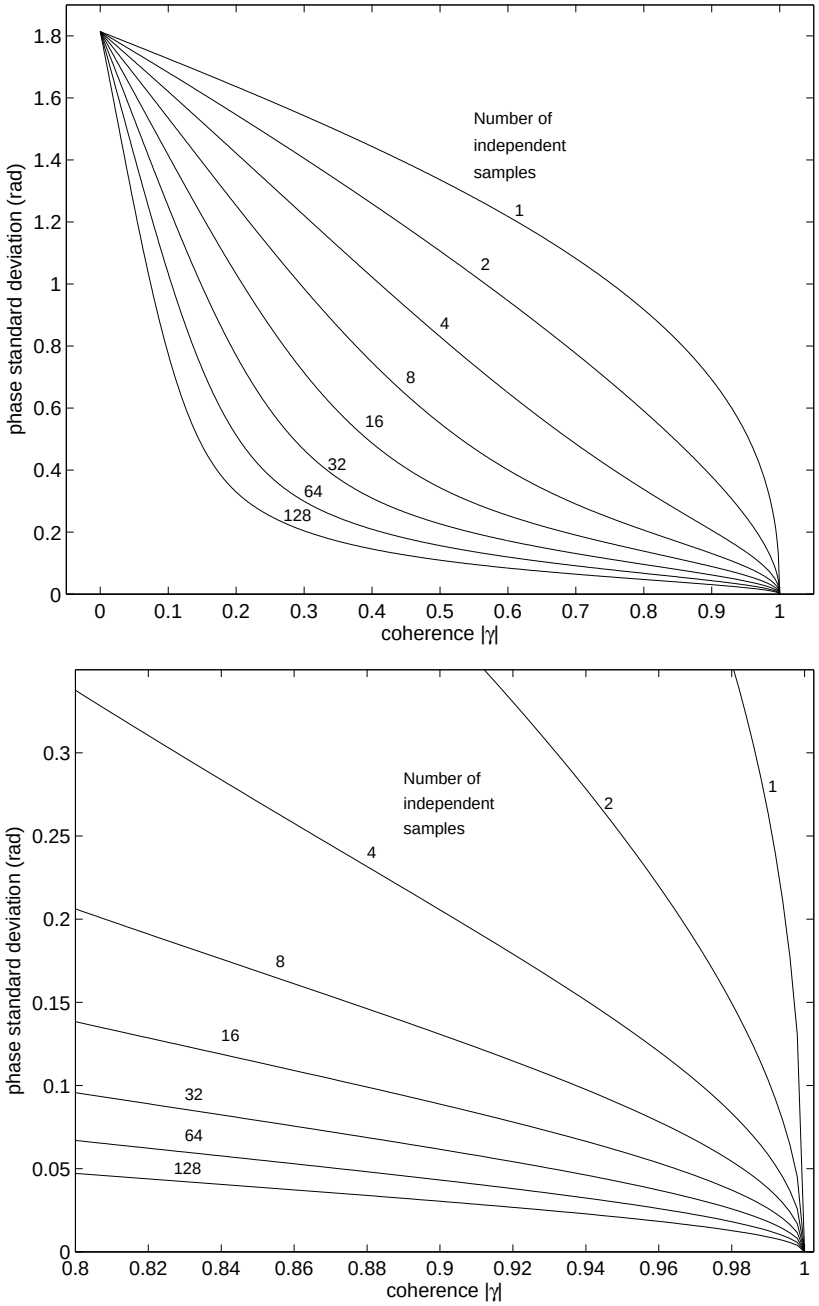


Figure 4.5 Phase difference standard deviation decreases as the number of independent samples (L) are combined. Upper plot shows the full coherence range 0.0–1.0, the lower plot showing the range 0.8–1.0, a more typical operating range for interferometry applications. For a given coherence level large gains are initially made by increasing from a single look. For higher numbers of looks the gains become smaller.

[Goodman, 1985, Lombardini, 1996, Lombardo and Lombardini, 1997]

$$p(\mathbf{d}|a,\phi) = \frac{\exp\left(-(\mathbf{d}-\bar{\mathbf{d}})^H \mathbf{K}_d^{-1}(\mathbf{d}-\bar{\mathbf{d}})\right)}{\pi^N \det\left(\mathbf{K}_d\right)}, \tag{4.22}$$

where $\mathbf{K_d}$ is the covariance matrix given by

$$\mathbf{K}_d = \mathrm{E}\left\{(\mathbf{d}-\bar{\mathbf{d}})(\mathbf{d}-\bar{\mathbf{d}})^H\right\}, \tag{4.23}$$

and $\mathbf{d}$ is the vector of data measurements d_h with a mean of $\bar{\mathbf{d}}$. Assuming the data signals are zero mean ($\bar{\mathbf{d}} = \mathbf{0}$) the likelihood function reduces to [Seymour and Cumming, 1994]

$$p(\mathbf{d}|a,\phi) = \frac{1}{\pi^N \det\left(\mathbf{K}_d\right)} \exp\left(-\mathbf{d}^H \mathbf{K}_d^{-1}\mathbf{d}\right), \tag{4.24}$$

with covariance matrix

$$\mathbf{K}_d = \mathrm{E}\left\{\mathbf{d}\mathbf{d}^H\right\}. \tag{4.25}$$

The maximum likelihood phase estimate ϕ_{ML} that maximises the likelihood is

$$\phi_{ML} = \arg\max_{\phi} p(\mathbf{d}|a,\phi). \tag{4.26}$$

This is equivalent to minimising the argument of the exponential given in (4.24),

$$\phi_{ML} = \arg\min_{\phi} C\left(\phi\right) \tag{4.27}$$

where the cost function $C\left(\phi\right)$ is

$$C\left(\phi\right) = -\mathbf{d}^H \mathbf{K_d}^{-1}\mathbf{d}. \tag{4.28}$$

4.4.1 Two receivers

For the case of two receivers, the received signals from (4.21) are

$$d_1 = a_1 m_1 + n_1, \tag{4.29a}$$

$$d_2 = a_2 m_2 \exp\left(j\phi\right) + n_2. \tag{4.29b}$$

Again, assuming the two signals are of equal amplitude ($a = a_1 = a_2$),

$$d_1 = a m_1 + n_1, \tag{4.30a}$$

$$d_2 = a m_2 \exp\left(j\phi\right) + n_2. \tag{4.30b}$$

Expanding $\mathbf{K}_d$ of (4.25) for the two receiver case (and assuming zero mean signals),

$$\mathbf{K}_d = \mathrm{E}\left\{ \begin{bmatrix} d_1 d_1^* & d_1 d_2^* \\ d_1^* d_2 & d_2 d_2^* \end{bmatrix} \right\} = \begin{bmatrix} \mathrm{E}\left\{|d_1|^2\right\} & \mathrm{E}\left\{d_1 d_2^*\right\} \\ \mathrm{E}\left\{d_1^* d_2\right\} & \mathrm{E}\left\{|d_2|^2\right\} \end{bmatrix}. \tag{4.31}$$

The main diagonal and cross terms can be expanded, assuming the speckle and additive noise are independent,

$$\begin{aligned} \mathrm{E}\left\{|d_1|^2\right\} &= \mathrm{E}\left\{|am_1 + n_1|^2\right\} \\ &= |a|^2 \mathrm{E}\left\{|m_1|^2\right\} + 2|a|\mathrm{E}\left\{m_1\right\}\mathrm{E}\left\{n_1\right\} + \mathrm{E}\left\{|n_1|^2\right\}, \qquad (4.32) \\ \mathrm{E}\left\{d_1 d_2^*\right\} &= \mathrm{E}\left\{(am_1 + n_1)\left(a^* m_2^* \exp\left(-j\phi\right) + n_2^*\right)\right\} \\ &= \mathrm{E}\left\{|a|^2 m_1 m_2^* \exp\left(-j\phi\right)\right\} + \mathrm{E}\left\{am_1 n_2^*\right\} \\ &\quad + \mathrm{E}\left\{a^* m_2^* n_1 \exp\left(-j\phi\right)\right\} + \mathrm{E}\left\{n_1 n_2^*\right\}. \qquad (4.33) \end{aligned}$$

Since the noise and speckle are independent, (4.32) and (4.33) simplify to

$$\begin{aligned} \mathrm{E}\left\{|d_1|^2\right\} &= |a|^2 \mathrm{E}\left\{|m_1|^2\right\} + \mathrm{E}\left\{|n_1|^2\right\}, \qquad (4.34) \\ \mathrm{E}\left\{d_1 d_2^*\right\} &= |a|^2 \exp\left(-j\phi\right)\mathrm{E}\left\{m_1 m_2^*\right\}. \qquad (4.35) \end{aligned}$$

By defining the signal power S and noise power N as

$$\begin{aligned} S &= |a|^2 \mathrm{E}\left\{|m_h|^2\right\}, \qquad (4.36) \\ N &= \mathrm{E}\left\{|n_h|^2\right\}, \qquad (4.37) \end{aligned}$$

the covariance matrix terms become

$$\begin{aligned} \mathrm{E}\left\{|d_1|^2\right\} &= S + N, \qquad (4.38) \\ \mathrm{E}\left\{d_1 d_2^*\right\} &= S\alpha_{1,2} \exp\left(-j\phi\right), \qquad (4.39) \end{aligned}$$

where $\alpha_{h,i}$ is the spatial correlation coefficient of the two signals,

$$\alpha_{h,i} = \frac{\mathrm{E}\left\{m_h m_i^*\right\}}{\mathrm{E}\left\{|m_h|^2\right\}}. \tag{4.40}$$

Since $\mathrm{E}\left\{|m_h|^2\right\}$ is defined to be unity,

$$\alpha_{h,i} = \mathrm{E}\left\{m_h m_i^*\right\}. \tag{4.41}$$

Substituting (4.38) and (4.39) into (4.31) gives

$$\mathbf{K_d} = \begin{bmatrix} S+N & S\alpha_{1,2}\exp(-j\phi) \\ S\alpha_{2,1}\exp(j\phi) & S+N \end{bmatrix}, \tag{4.42}$$

$$= \sigma^2 \begin{bmatrix} 1 & \mu\exp(-j\phi) \\ \mu^*\exp(j\phi) & 1 \end{bmatrix}, \tag{4.43}$$

where σ^2 and μ are

$$\sigma^2 = S+N, \tag{4.44}$$

$$\mu = \frac{\alpha_{h,i}}{1+N/S} \tag{4.45}$$

$$= \frac{\alpha_{h,i}}{1+\text{SNR}^{-1}}, \tag{4.46}$$

and SNR is the signal to noise ratio (SNR $= S/N$). Often the correlation coefficient μ is assumed to be real, implying the spatial spectra of the the speckle components are symmetric around zero frequency [Gini et al., 2002]. This assumption simplifies the covariance matrix of (4.43) to

$$\mathbf{K_d} = \sigma^2 \begin{bmatrix} 1 & \mu\exp(-j\phi) \\ \mu\exp(j\phi) & 1 \end{bmatrix}. \tag{4.47}$$

The $\mathbf{K_d}$ matrix of (4.47) can be easily inverted, and substituted into the original likelihood function of (4.24),

$$C(\phi) = \frac{1}{\sigma^2(1-|\mu|^2)} [d_1^* d_2^*] \begin{bmatrix} 1 & -\mu^*\exp(j\phi) \\ -\mu^*\exp(-j\phi) & 1 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \tag{4.48}$$

$$= \frac{1}{\sigma^2(1-\mu^2)} \left[|d_1|^2 - \mu\exp(j\phi)\, d_1^* d_2 + |d_2|^2 - \mu^*\exp(-j\phi)\, d_1 d_2^*\right]. \tag{4.49}$$

Removing terms independent of ϕ,

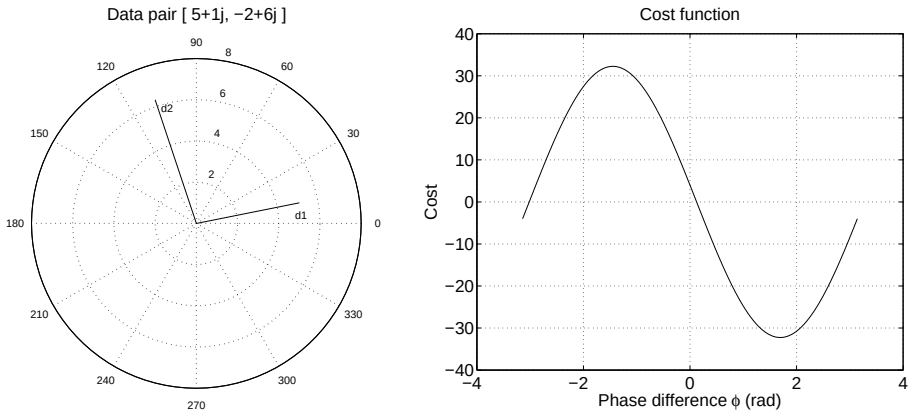
$$C(\phi) = -\exp(j\phi)\, d_1^* d_2 - \exp(-j\phi)\, d_1 d_2^* \tag{4.50}$$

$$= -2\mathcal{R}\left\{\exp(j\phi)\, d_1^* d_2\right\}. \tag{4.51}$$

The maximum likelihood phase difference ϕ_{ML} of Equation 4.27 can be found from inspection

$$\phi_{ML} = \arg(d_1 d_2^*). \tag{4.52}$$

Thus the maximum likelihood phase difference with two receivers is simply the smallest phase difference between them, consistent with (4.2). An example cost function for a two receiver case is shown in Figure 4.6.



(a) (b)

Figure 4.6 Example of the maximum likelihood phase difference cost function (b) for the data pair $\mathbf{d} = [5 + 1j, -2 + 6j]$ as shown in (a). The ML estimate cost function is a minimum at $\phi = 1.7$.

4.4.2 Three receivers

For the case of three evenly separated receivers with signal amplitudes assumed equal, the received signals using (4.21) are

$$d_1 = am_1 + n_1, \tag{4.53a}$$

$$d_2 = am_2 \exp(j\phi/2) + n_2, \tag{4.53b}$$

$$d_3 = am_3 \exp(j\phi) + n_3, \tag{4.53c}$$

where ϕ is defined as the phase difference of the outer two receivers and is the parameter to be estimated. Expanding $\mathbf{K}_d$ gives

$$\mathbf{K}_d = \mathrm{E}\left\{ \begin{bmatrix} d_1 d_1^* & d_1 d_2^* & d_1 d_3^* \\ d_1^* d_2 & d_2 d_2^* & d_2 d_3^* \\ d_1^* d_3 & d_2^* d_3 & d_3 d_3^* \end{bmatrix} \right\} = \begin{bmatrix} \mathrm{E}\left\{|d_1|^2\right\} & \mathrm{E}\left\{d_1 d_2^*\right\} & \mathrm{E}\left\{d_1 d_3^*\right\} \\ \mathrm{E}\left\{d_1^* d_2\right\} & \mathrm{E}\left\{|d_2|^2\right\} & \mathrm{E}\left\{d_2 d_3^*\right\} \\ \mathrm{E}\left\{d_1^* d_3\right\} & \mathrm{E}\left\{d_2^* d_3\right\} & \mathrm{E}\left\{|d_3|^2\right\} \end{bmatrix}. \tag{4.54}$$

Expanding these terms as shown for the two receiver case yields

$$\mathrm{E}\left\{|d_1|^2\right\} = S + N, \tag{4.55}$$

$$\mathrm{E}\left\{d_1 d_2^*\right\} = S\alpha_{1,2} \exp(-j\phi/2), \tag{4.56}$$

$$\mathrm{E}\left\{d_1 d_3^*\right\} = S\alpha_{1,3} \exp(-j\phi), \tag{4.57}$$

with a resulting $\mathbf{K_d}$ matrix

$$\mathbf{K_d} = \begin{bmatrix} S+N & S\alpha_{1,2}\exp(-j\phi/2) & S\alpha_{1,3}\exp(-j\phi) \\ S\alpha_{2,1}\exp(j\phi/2) & S+N & S\alpha_{2,3}\exp(-j\phi/2) \\ S\alpha_{3,1}\exp(j\phi) & S\alpha_{3,2}\exp(j\phi/2) & S+N \end{bmatrix}, \quad (4.58)$$

$$= \sigma^2 \begin{bmatrix} 1 & \mu_{1,2}\exp(-j\phi/2) & \mu_{1,3}\exp(-j\phi) \\ \mu^*_{2,1}\exp(j\phi/2) & 1 & \mu_{2,3}\exp(-j\phi) \\ \mu^*_{3,1}\exp(j\phi) & \mu^*_{3,2}\exp(j\phi/2) & 1 \end{bmatrix}. \quad (4.59)$$

where σ^2 and $\mu_{h,i}$ are

$$\sigma^2 = S+N, \quad (4.60)$$

$$\mu_{h,i} = \frac{\alpha_{h,i}}{1+\text{SNR}^{-1}}. \quad (4.61)$$

After finding the inverse of $\mathbf{K_d}$, and substitution into the cost function of (4.28),

$$C = \frac{1}{\sigma^2\left(1-\mu^*_{1,2}\mu_{1,2}-\mu^*_{2,3}\mu_{2,3}-\mu^*_{1,3}\mu_{1,3}+\mu_{1,2}\mu_{2,3}\mu^*_{1,3}+\mu^*_{1,2}\mu^*_{2,3}\mu_{1,3}\right)}\times$$

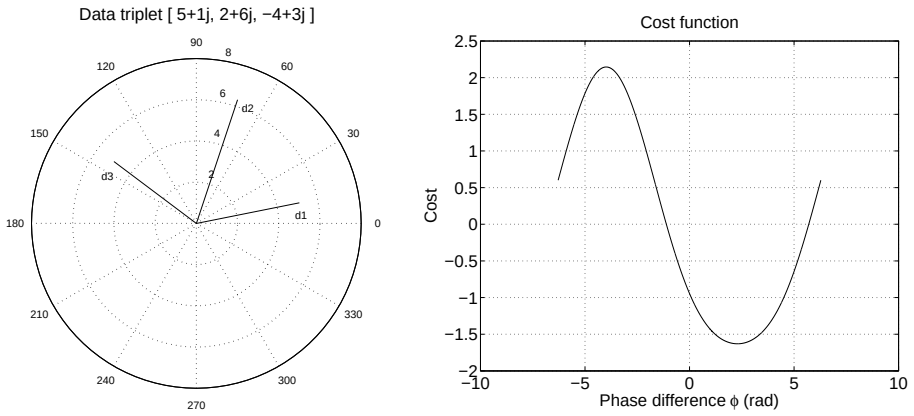
$$\begin{pmatrix} d^*_1d_1 - d^*_1d_1\mu_{2,3}\mu^*_{2,3} - d^*_1d_2\mu_{1,2}\exp(-j\phi/2) + d^*_1d_2\mu_{1,3}\mu^*_{2,3}\exp(-j\phi/2) \\ +d^*_1d_3\mu_{1,2}\mu_{2,3}\exp(-j\phi) - d^*_1d_3\mu_{1,3}\exp(-j\phi) + d^*_2d_1\mu^*_{1,3}\mu_{2,3}\exp(j\phi/2) \\ -d^*_2d_1\mu^*_{1,2}\exp(j\phi/2) + d^*_2d_2 - d^*_2d_2\mu^*_{1,3}\mu_{1,3} - d^*_2d_3\mu_{2,3}\exp(-j\phi/2) \\ +d^*_2d_3\mu_{1,3}\mu^*_{1,2}\exp(-j\phi/2) + d^*_3d_1\mu^*_{1,2}\mu^*_{2,3}\exp(j\phi) - d^*_3d_1\mu^*_{1,3}\exp(j\phi) \\ -d^*_3d_2\mu^*_{2,3}\exp(j\phi/2) + d^*_3d_2\mu^*_{1,3}\mu_{1,2}\exp(j\phi/2) + d^*_3d_3 + d^*_3d_3\mu^*_{1,2}\mu_{1,2} \end{pmatrix}$$

$$= \frac{1}{\sigma^2\left(1-\mu^*_{1,2}\mu_{1,2}-\mu^*_{2,3}\mu_{2,3}-\mu^*_{1,3}\mu_{1,3}+\mu_{1,2}\mu_{2,3}\mu^*_{1,3}+\mu^*_{1,2}\mu^*_{2,3}\mu_{1,3}\right)}\times$$

$$\begin{pmatrix} d^*_1d_2\exp(-j\phi/2)\left(\mu_{1,3}\mu^*_{2,3}-\mu_{1,2}\right) \\ +d_1d^*_2\exp(j\phi/2)\left(\mu^*_{1,3}\mu_{2,3}-\mu^*_{1,2}\right) \\ +d^*_2d_3\exp(-j\phi/2)\left(\mu_{1,3}\mu^*_{1,2}-\mu_{2,3}\right) \\ +d_2d^*_3\exp(j\phi/2)\left(\mu^*_{1,3}\mu_{1,2}-\mu^*_{2,3}\right) \\ +d^*_1d_3\exp(-j\phi)\left(\mu_{1,2}\mu_{2,3}-\mu_{1,3}\right) \\ +d_1d^*_3\exp(j\phi)\left(\mu^*_{1,2}\mu^*_{2,3}-\mu^*_{1,3}\right) \end{pmatrix}. \quad (4.62)$$

Removing terms independent of ϕ and combining conjugate pairs using the rule $a+a^* = 2\mathcal{R}\{a\}$, the cost function of (4.62) reduces to

$$\begin{aligned} C(\phi) =& \mathcal{R}\left\{d^*_1d_2\exp(-j\phi/2)\left(\mu_{1,3}\mu^*_{2,3}-\mu_{1,2}\right)\right\} \\ &+ \mathcal{R}\left\{d^*_2d_3\exp(-j\phi/2)\left(\mu_{1,3}\mu^*_{1,2}-\mu_{2,3}\right)\right\} \\ &+ \mathcal{R}\left\{d^*_1d_3\exp(-j\phi)\left(\mu_{1,2}\mu_{2,3}-\mu_{1,3}\right)\right\}. \end{aligned} \quad (4.63)$$



(a) (b)

Figure 4.7 Example of the maximum likelihood phase difference cost function (b) for the data triplet $\mathbf{d} = [5 + 1j, 2 + 6j, -4 + 3j]$ as shown in (a). The ML estimate cost function is a minimum at $\phi = 2.3$.

Often the correlation coefficients $\mu_{h,i}$ are assumed to be real, implying the spatial spectra of the the speckle components are symmetric around zero spatial frequency [Gini et al., 2002]. This assumption simplifies the cost function of (4.63) to

$$\begin{aligned} C(\phi) = &(\mu_{1,3}\mu_{2,3} - \mu_{1,2})\,\mathcal{R}\left\{d_1^* d_2 \exp(-j\phi/2)\right\} \\ &+ (\mu_{1,3}\mu_{1,2} - \mu_{2,3})\,\mathcal{R}\left\{d_2^* d_3 \exp(-j\phi/2)\right\} \\ &+ (\mu_{1,2}\mu_{2,3} - \mu_{1,3})\,\mathcal{R}\left\{d_1^* d_3 \exp(-j\phi)\right\}. \end{aligned} \tag{4.64}$$

An example of a three element cost function is shown in Figure 4.7.

4.4.3 Multiple frequency bands

Multiple frequency bands can also be employed to extend the unambiguous phase interval, and to reduce the variance of the estimated phase difference. With multiple frequency bands, the time delay measured between hydrophones is frequency independent, assuming the speed of sound is also independent of frequency. This now gives a system model of

$$d_{h,k} = a_{h,k} m_{h,k} \exp(j2\pi b_{h,1} f_k \Delta\tau) + n_{h,k}, \tag{4.65}$$

for N_h hydrophones and N_k frequency bands. Here, $b_{h,1}$ is the normalised baseline between hydrophones h and 1. With this model, the covariance matrix $\mathbf{K}_{d,k}$ for each frequency band has elements of

$$\alpha_{i,j,k} = \sigma_k^2 \mu_{i,j,k} \exp(-j2\pi b_{i,j} f_k \Delta\tau), \tag{4.66}$$

where $\mu_{j,i,k} = \mu_{i,j,k}^*$ and $\mu_{i,i,k} = 1$. By assuming all the frequency bands are independent, an overall covariance matrix $\mathbf{K}_{\mathbf{d},\mathbf{k}}$ can be formed from the matrices of each

frequency band via diagonal concatenation on the main diagonal, with zeros in the outer quadrants. For example, with a three receiver ($N_h = 3$), two frequency band ($N_k = 2$) system such as the KiwiSAS-III system, $\mathbf{K_d}$ becomes

$$\mathbf{K_d} = \begin{bmatrix} \mathbf{K_{d,1}} & 0 \\ 0 & \mathbf{K_{d,2}} \end{bmatrix}, \tag{4.67}$$

where $K_{d,k}$ are each 3×3 matrices. Since the frequency bands are assumed to be independent, this is equivalent to forming an overall probability distribution as the product of the probabilities of each frequency band,

$$p(\mathbf{d}|a, \Delta t) = \prod_{k=1}^{N_k} \frac{1}{\pi^N \det\left(\mathbf{K}_{d,k}\right)} \exp\left(-\mathbf{d}_k^H \mathbf{K}_{d,k}^{-1} \mathbf{d}_k\right). \tag{4.68}$$

Similarly, this independence assumption allows an overall cost function to be formed as the sum of each frequency bands cost function,

$$C\left(\Delta t\right) = \sum_{k=1}^{N_k} C_k\left(\Delta t\right). \tag{4.69}$$

For the three hydrophone case,

$$\begin{aligned} &C_k\left(\Delta\tau\right) = \\ &\mathcal{R}\left\{d_{1,k}^* d_{2,k} \exp\left(-j2\pi f_k b_{2,1} \Delta\tau\right)\left(\mu_{1,3}\mu_{2,3}^* - \mu_{1,2}\right)\right\} \\ &+ \mathcal{R}\left\{d_{2,k}^* d_{3,k} \exp\left(-j2\pi f_k b_{3,2} \Delta\tau\right)\left(\mu_{1,3}\mu_{1,2}^* - \mu_{2,3}\right)\right\} \\ &+ \mathcal{R}\left\{d_{1,k}^* d_{3,k} \exp\left(-j2\pi f_k b_{3,1} \Delta\tau\right)\left(\mu_{1,2}\mu_{2,3} - \mu_{1,3}\right)\right\}, \end{aligned} \tag{4.70}$$

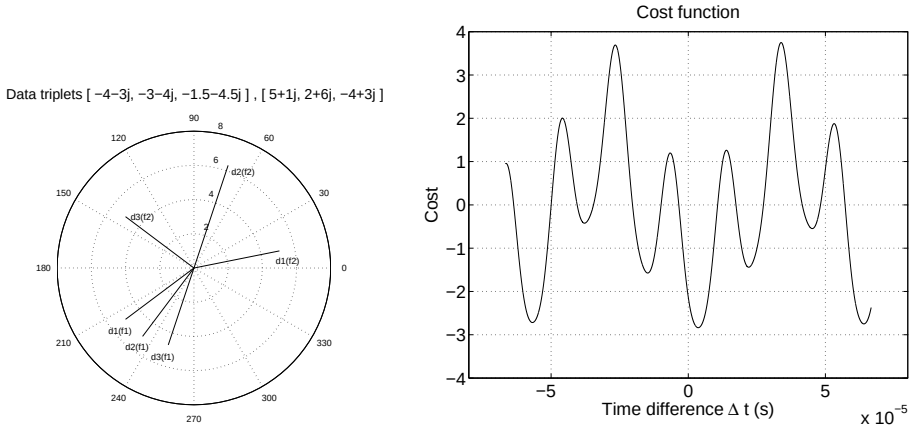
assuming $\mu_{i,j,k} = \mu_{i,j}$ for notational convenience, although in practice the critical baseline and thus coherence varies with frequency. The maximum likelihood phase estimate is thus found using

$$\widehat{\Delta\tau}_{\mathrm{ML}} = \arg\min_{\Delta\tau} \sum_{k=1}^{N_k} C_k\left(\Delta\tau\right). \tag{4.71}$$

An example typical cost function is shown in Figure 4.8 for the $N_h = 3$, $N_k = 2$ case.

4.4.4 Combining multiple along-track looks

Multiple along-track looks can be combined in a ML sense by simply summing the cost functions of each of the individual along-track looks. The cost function for the three



(a) (b)

Figure 4.8 Example of the maximum likelihood time difference cost function (b) for the two data triplets $\mathbf{d} = [-4-3j, -3-4j, -1.5-4.5j], [5+1j, 2+6j, -4-3j]$ with centre frequencies of f_c=30 kHz and 100 kHz respectively, as shown in (a). The ML estimate cost function is a minimum at $\Delta t = 3.63 \times 10^{-6}$.

receiver system with N_p along-track looks is,

$$\begin{aligned} C = &\left(\mu_{1,3}\mu_{2,3} - \mu_{1,2}\right) \mathcal{R}\left\{\sum_{i=1}^{N_p} d_{1,i}^{*} d_{2,i} \exp\left(-j\phi\left(1-p\right)\right)\right\} \\ &+ \left(\mu_{1,3}\mu_{1,2} - \mu_{2,3}\right) \mathcal{R}\left\{\sum_{i=1}^{N_p} d_{2,i}^{*} d_{3,i} \exp\left(-j\phi p\right)\right\} \\ &+ \left(\mu_{1,2}\mu_{2,3} - \mu_{1,3}\right) \mathcal{R}\left\{\sum_{i=1}^{N_p} d_{1,i}^{*} d_{3,i} \exp\left(-j\phi\right)\right\}. \end{aligned} \tag{4.72}$$

4.4.5 Searching the Cost Function

When solving problems using a maximum likelihood approach, the final step in finding a single solution is to search the resulting cost function to find the minimum. Numerical search algorithms are often slow to compute, a large number of iterations are often required to converge on an answer. Therefore, any method available to reduce the number of iterations needed is advantageous. Without any additional *a priori* information or constraints, the search must be carried out over the entire solution space and is impossible to solve for many cost functions. The introduction of prior information can greatly reduce the parameter search space, especially if this information includes fixed constraints on the solution space. For InSAS, fixed limits can be introduced to the parameter search space based on prior information of the range of seafloor heights expected within the scene.

For the cost functions derived in Sections 4.4.1-4.4.4, the cost function is the sum of

several sinusoidal functions with known frequencies. This allows the cost function to be coarsely searched using a fixed spacing of test points, from which a smaller constrained area can be searched to find the global minimum. The minimum spacing required to coarsely estimate the minimum is determined by the maximum rate of change of the cost function, dictated by the highest frequency component in the cost function. In the case of a flat, vertical receiver array, this component is that of the outermost pair of receivers, at the highest operating frequency. From this course sampling a narrower, bounded search can be performed to find the true minimum of the cost function, giving the overall maximum likelihood estimate. Due to numerical precision errors, the initial rate of sampling should be further increased to higher than the Nyquist rate, ensuring the global minima is found. An example of this search strategy for a typical InSAS cost function is shown in Figure 4.9. First, limits are placed on the cost function search. Within these bounds a coarse search is performed at sample spacings of $\frac{1}{2f_{\text{cmax}}}$. Finally, the minimum cost is found from the minimum coarse cost function evaluation. For the bathymetric reconstructions in this thesis, the MATLAB[1] computer package was used. Specifically, the function `fminbnd` was used for searching the cost function once the region containing the global minimum was found.

4.5 INITIAL GROUNDPLANE DATA CORRECTION

As shown in Section 3.4, footprint mis-alignment is the normally the largest loss of coherence for an uncorrected InSAS system. The effect of footprint mis-alignment can be reduced significantly by re-sampling the echo responses from each receiver of the interferometric array to a common ground-plane of height approximately equal to the true seafloor surface. By incorporating the geometry of the receiver array during this re-sampling, the mis-alignment between the receivers can be reduced. For each receiver R_n the geometry is as shown in Figure 4.10, where s is the slant-range data, and g the remapped ground-plane data. The envelope remapping is then given by the equations:

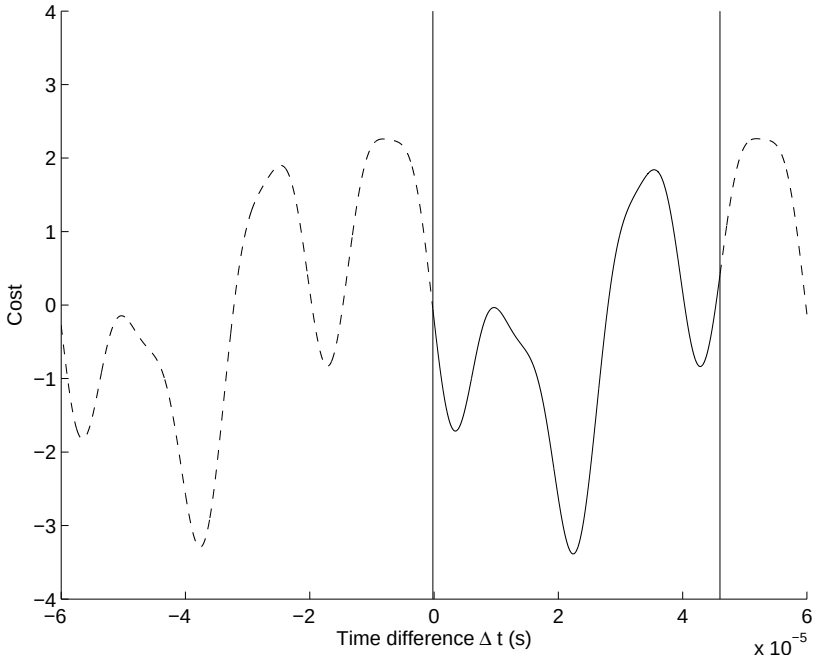
$$\begin{aligned} s_n(x) &= \sqrt{g(x)^2 + (g(z) - P_n(z))^2}, \\ s_n(y) &= g(y). \end{aligned} \tag{4.73}$$

The phase correction term required to compensate for the demodulation of the original acoustic signal is then

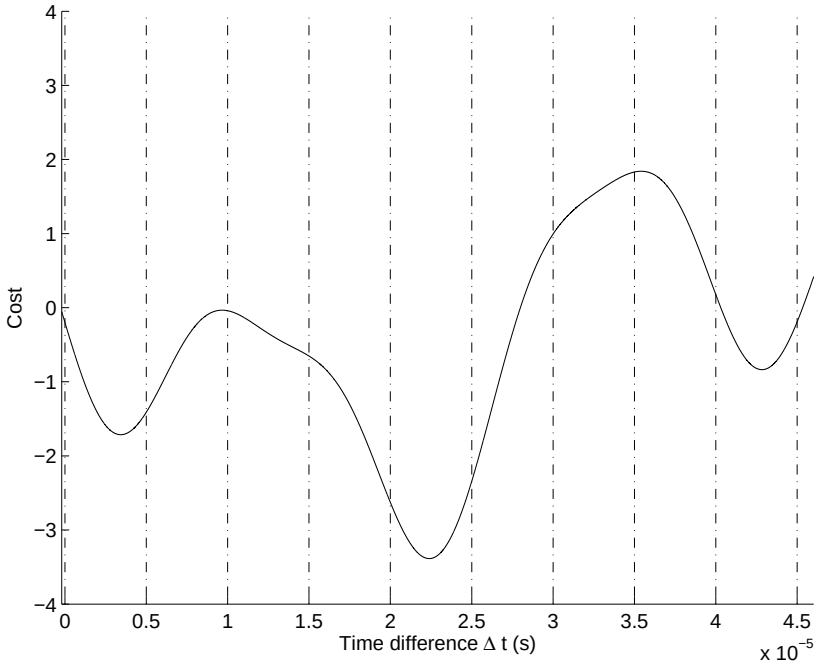
$$\varphi = \exp\left(\frac{-j4\pi f_c s(x)}{c}\right). \tag{4.74}$$

The overall effect of this process can be seen in Figure 4.11, for the case of a single point reflector located 45 m directly across-track, 5 m below the sonar. Notice the peaks of the original slant range curves are centred at approximately 45.2 m, and

[1] www.mathworks.com



(a) Overall cost with coarse search limits shown.



(b) Search grid within coarse limits.

Figure 4.9 Cost function search strategy. (a) Initial bounds (vertical solid lines) are calculated for the search range based on the expected height of the seafloor. Within these bounds, a initial search is performed at spacings of $t = \frac{1}{2f_{\mathrm{cmax}}}$ (vertical dashed lines of (b)). The minimum of these initial searches (t=2×10^{-5}) then gives a starting value for a minimum search, eliminating the chance of finding a local minima such as at t=0.3×10^{-5}.

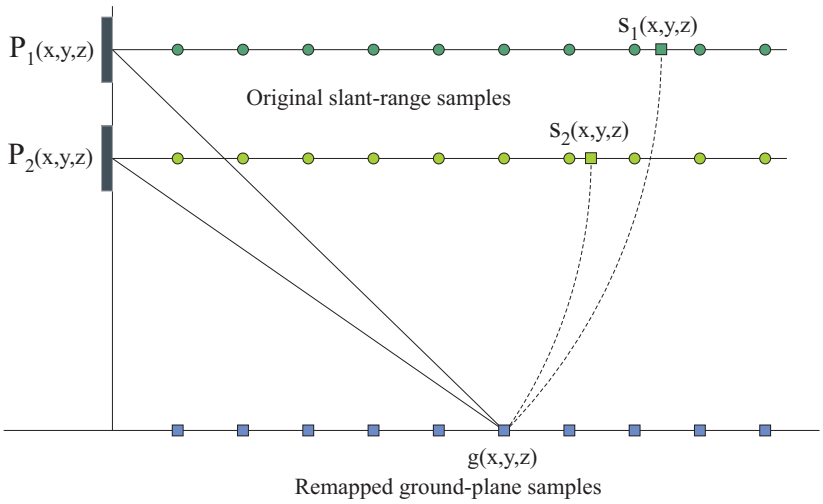


Figure 4.10 Remapping slant-range plane to ground-plane. The transformation is made by rotating each desired ground-plane sample position around each phase centre. The intersection of the rotation and the slant-range plane gives the point of interpolation from the slant-plane data.

after remapping to the correct groundplane they are centred on the original simulation position of 45 m. For a flat seafloor scene, ground-plane remapping can be seen by comparing the interferograms of the original slant-range plane data, and after remapping to the 'correct' ground-plane as shown in Figure 4.12. The original interferogram is very speckly in nature, and shows several interference fringes in range across the image. After remapping to the 'correct' height-plane the speckle has been reduced, and the interference fringes have been removed leaving a mean phase difference of zero. The speckle reduction is as a result of increasing the correlation between the scenes with corresponding pixels footprints now aligned. The remaining speckle is from the slightly different angle of viewing of the scene causing baseline decorrelation. Bands of higher speckle can be seen across the interferogram at ranges of 23 m, 11 m, 6 m, and several < 5 m corresponding to the nulls in the vertical beam pattern of the transducers.

4.5.1 Belief Propagation

One technique to make an initial estimate of the seafloor shape is to use belief propagation. Belief propagation is an efficient technique based on local message passing, used to find the most likely state of a system given any available evidence. By appropriately modeling the system, this approach can be applied to bathymetric data to produce an improved estimate the sea floor. Although originally proposed by [Pearl, 1988] for performing probabilistic reasoning on Bayesian networks, the technique has since been successfully applied to the stereo imaging field [Sun et al., 2002, Petrovic et al., 2001, Minagawa et al., 2002]. In most of this work, the system is modeled as a hidden Markov Random Field (MRF). Belief propagation is then applied to this model,

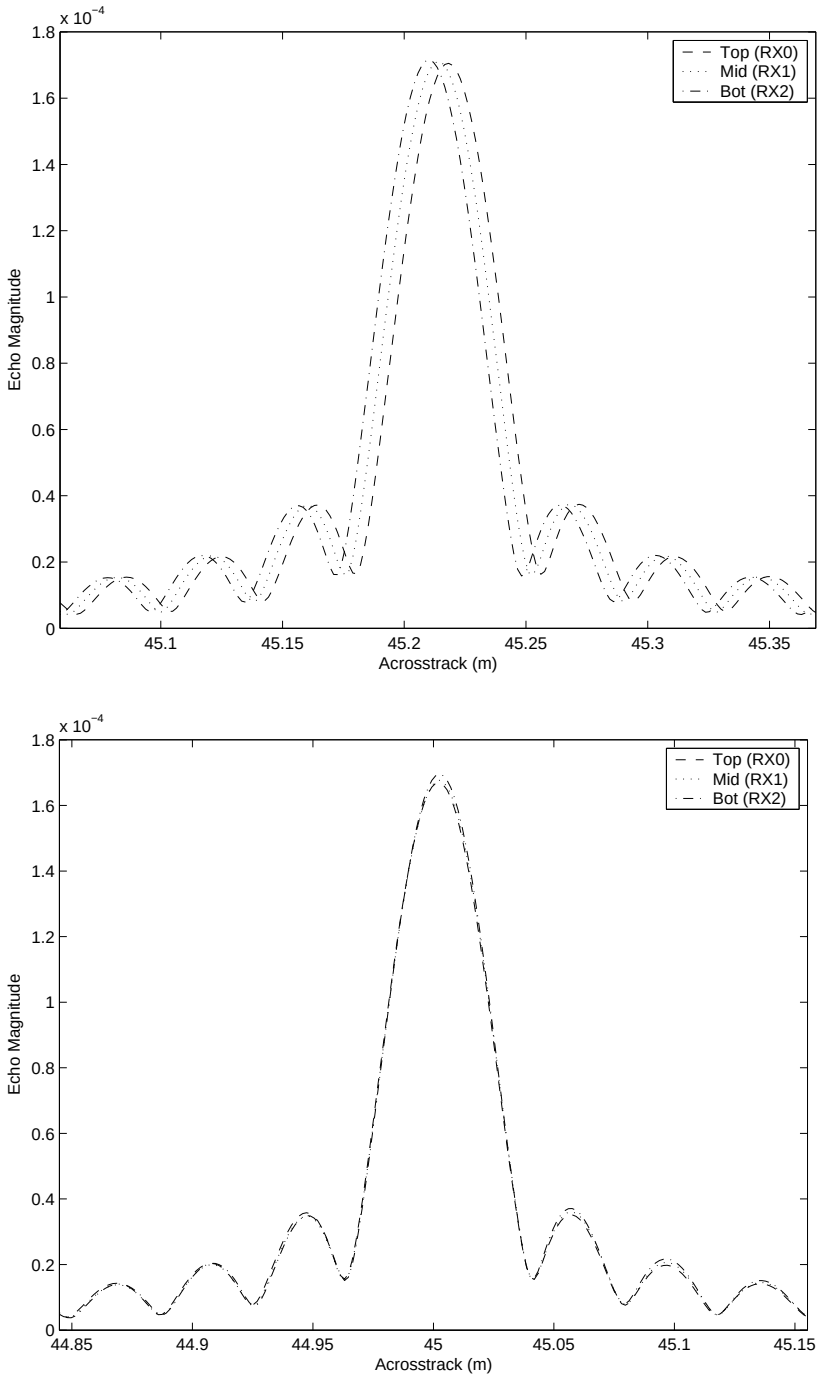


Figure 4.11 Signal envelope of a single point target at 40 m range, 5 m below sonar. (a) reconstructed on slant range plane, and (b) remapped to correct groundplane.

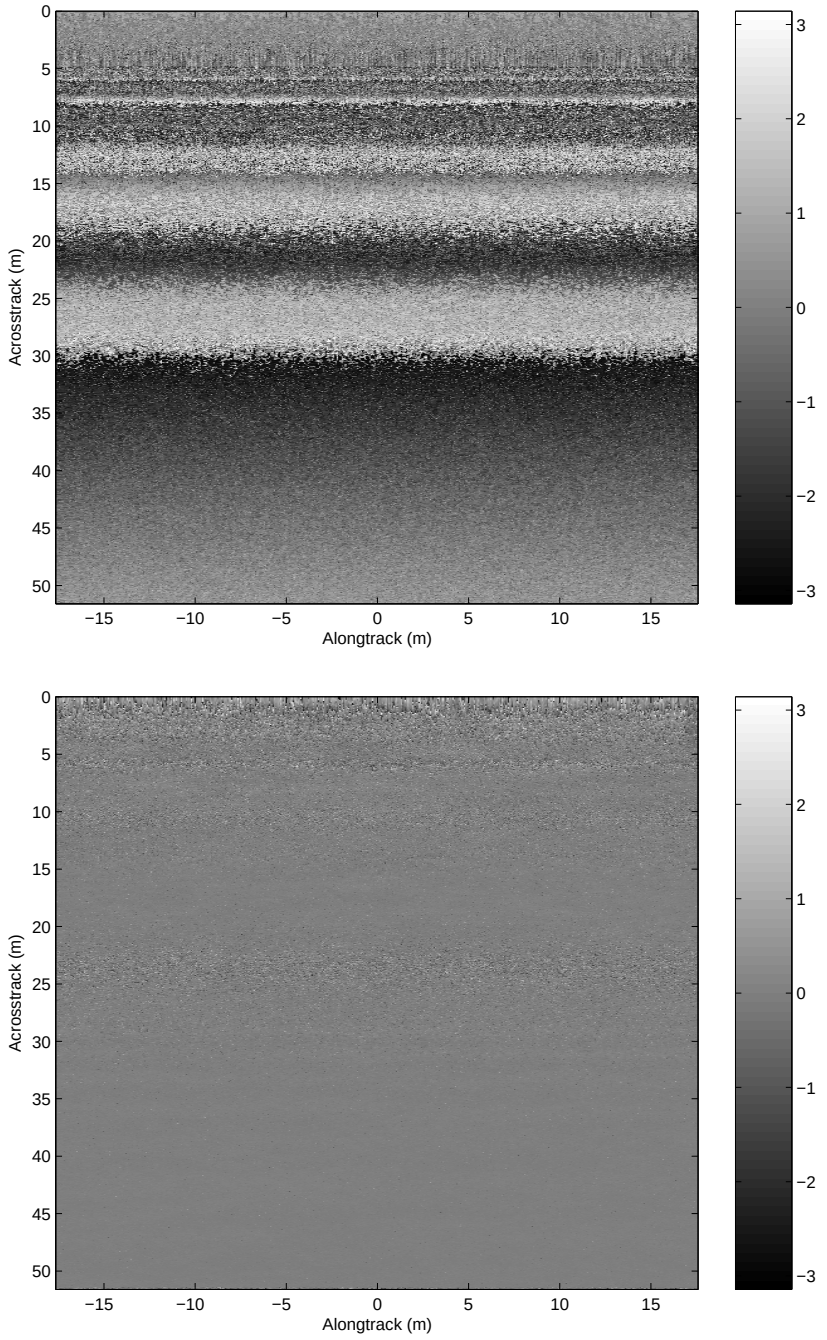


Figure 4.12 Interferogram for (a) slant-range plane data, and (b) ground-plane data remapped to the 'correct' height. Receiver separation 0.15 m.

acting as a statistical filter to find the most likely surface, given a volume of measured likelihood estimates and the expected variation between neighbouring points.

To apply this technique to bathymetric InSAS data, the scene is represented as a connected 2D array of nodes, whose value corresponds to surface height. Associated with each node are a set of beliefs, each give a measure how likely the surface is to be at any particular height. The objective is to calculate these beliefs as accurately as possible, by propagating the original data throughout the network. More details of belief propagation are presented in Appendix A. To prepare the data for surface fitting with belief propagation, first the data must be re-sampled onto a volume of sea-floor height planes. From the re-sampled volume data a metric of probability is generated, through which the belief propagation algorithm fits a surface. The implication of using quantised height planes in the generation of the probability volume is the final heightmap estimation is also quantised to the same planes. This effect can be seen in the results of Appendix A, where the continuous surface of the structure is quantised to discrete heights in the final height estimation. The quantisation can be reduced by using more discrete heights in the probability volume generation, however the computational requirements become impractically large. Instead, belief propagation is more suited to generating an initial estimate of the seafloor surface height. This initial estimate can then be used to correct the original data, reducing the coherence loss of footprint mis-alignment. After performing this initial correction, the final height estimate can be refined using maximum likelihood techniques.

4.6 EXPANDING THE UNAMBIGUOUS PHASE INTERVAL

In its simplest form, an interferometer consists of two receivers separated by some distance D. After ideal correction for all coherence loss factors listed in Chapter 3, the coherence between these receivers approaches $\gamma = 1.0$. However, even in this ideal situation, the unambiguous range interval (URI) is still limited to an interval of 2π since phase can only be measured modulo 2π. This limits the usable baseline separation D to less than a few wavelengths. Beyond this limit, phase unwrapping techniques [Goldstein et al., 1988, Banks, 2002, Bonifant, 1999] must be used to retrieve the absolute phase between the receivers. Several publications in the SAR field have explored the use of multiple receivers [Lombardini and Lombardo, 1996, Massonnet et al., 1996, Corsini et al., 1999, Lombardini, 1996, Xu et al., 1994] and multiple frequency bands [Xu et al., 1994, Lanari et al., 1996] to extend the URI.

4.6.1 Interferometric array vertical spacing

The ML estimator of (4.28) has an unambiguous phase range for the estimation of ϕ limited by the number of receivers used. For example, the two receiver case has an

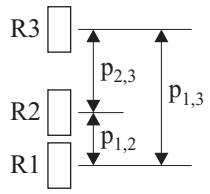


Figure 4.13 Unequally spaced receiver array. Separation constant defined as $p = p_{23}/b$.

unambiguous phase range of $\pm\pi$, and the three (evenly spaced) receiver case a range of $\pm 2\pi$. This range can be extended by using of an array of unequally separated receivers, forming multiple interferometric pairs with differing baselines [Corsini et al., 1997, Berizzi et al., 1997]. This technique is sometimes referred to as vernier interferometry [Wilby, 1999]. The separation constant of an unequally spaced, three-receiver array can be defined as the ratio of the upper pair of receivers to the total spacing of the array, as shown in Figure 4.13. Defining this separation constant,

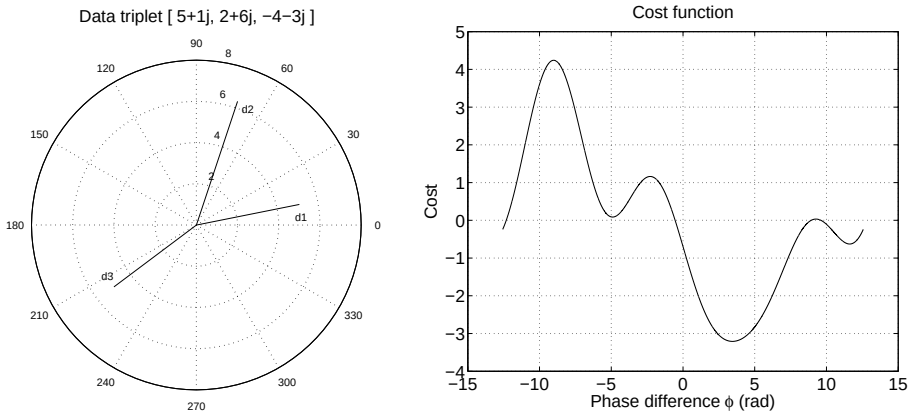
$$p = \frac{p_{2,3}}{p_{1,3}}, \tag{4.75}$$

the ML cost function to be minimised becomes,

$$\begin{aligned} C(\phi) = &(\mu_{1,3}\mu_{2,3} - \mu_{1,2})\,\mathcal{R}\left\{d_1^* d_2 \exp\left(-j\phi\,(1-p)\right)\right\} \\ &+ (\mu_{1,3}\mu_{1,2} - \mu_{2,3})\,\mathcal{R}\left\{d_2^* d_3 \exp\left(-j\phi p\right)\right\} \\ &+ (\mu_{1,2}\mu_{2,3} - \mu_{1,3})\,\mathcal{R}\left\{d_1^* d_3 \exp\left(-j\phi\right)\right\}. \end{aligned} \tag{4.76}$$

If $p = 1/n$, where n is a natural number, the unambiguous phase range now becomes $(-n\pi, n\pi]$. However, this increase in unambiguous phase range also requires the coherence term $\mu_{2,3}$ to remain sufficiently high as the separation $p_{2,3}$ is increased. Typically the baseline of an InSAS system is only a few wavelengths is length, limiting the maximum phase interval. Therefore, the factor of increase of the URI required for InSAS geometries is often low. An example of a typical cost function for an unevenly spaced, three-receiver, interferometric array is shown in Figure 4.14.

The coherence between each of the receiver pairs will decrease as the spacing between the receiver pair is increased.



(a) (b)

Figure 4.14 Example of the maximum likelihood phase difference cost function (b) for the data triplet $\mathbf{d} = [5 + 1j, 2 + 6j, -4 - 3j]$ as shown in (a) with an unequally spaced receiver set ($p = 2/3$). The ML estimate cost function is a minimum at $\phi = 3.4$.

Chapter 5

INTERFEROMETRIC SAS DESIGN AND PERFORMANCE

For a practical InSAS system a large set of parameters need to be chosen for the system design. Many of these parameters are governed either by desired imaging performance, or by practical construction limitations, often in conflict with each other. For the design discussion presented here, the practical considerations of constructing the platform are largely ignored, and instead an 'ideal' sonar platform considered. However, where applicable, comments are given for cases where construction would be impractical.

It is important to note the differences in resolution definitions in each of the three dimensions. Traditionally, the term resolution refers to the resolving power of the system, the minimum distance between two points in the scene able to be separated. This is the definition used in a sonar system for the across- and along-track resolution limits. For the case of height, resolution is poor, since two vertically separated points in the scene cannot be distinguished in the height map [Hayes et al., 2005]. Moreover, for a single interferometric pair of receivers, resolution is totally meaningless since only one angle of arrival can be estimated. Rather, in the z-direction the correct term is that of precision rather than resolution. The height estimation precision is defined as the error between the estimate and the true value for a single point of the image [Born and Wolf, 1999].

The implication of poor vertical resolution is the need for an assumption of only one point of reflection within each resolution cell, allowing only one angle-of-arrival at the receiver array to be estimated. If there is more than one angle-of-arrival from within a single resolution cell, this assumption is violated. Multiple angles-of-arrival can occur in situations of layover (see Section 5.8), or multipath reflections from the sea-surface (see Section 5.7). Some attempts to estimate multiple angles-of-arrival have been made using super-resolution techniques, covered in more detail in Section 5.7.1.

5.1 ACROSS-TRACK RESOLUTION DESIGN IMPLICATIONS

The theoretical resolution for a SAS system in the across-track direction, δ_x, is determined by the imaging bandwidth, B,

$$\delta_x = \frac{c}{2B}, \tag{5.1}$$

where c is the speed of sound in water, and B is the largest continuous segment of frequency used by the system [Hawkins, 1996]. The across-track resolution is independent of frequency, determined solely by the signal bandwidth. However, practical limits exist on the frequency spectrum able to be used since absorption in water increases with frequency [Urick, 1975]. Here it is also assumed the imaging is performed on a plane horizontal to the sonar platform such that the ground plane resolution δ_x is equal to the slant plane resolution δ_r. This is a reasonable approximation for the angles of imaging typical in a shallow water environment. The resolution given by (5.1) is, however, a theoretical limit requiring noise-free conditions, thus providing perfect matched filtering of the transmitted signal. Often, this is not the case in a real system because of noise in the acoustic channel, and distortion from the sonar transmitter and hydrophone arrays. Also, a spectral window is often applied to reduce the sidelobe height of the matched filtering, with an effect of lowering the effective bandwidth of the system (see Section 3.5.4). Therefore, to obtain a desired resolution a greater bandwidth than the theoretical limit of (5.1) is often required.

As part of the height estimation process, it is desirable to have several independent looks of the same scene, explored in more detail in Section 5.3. One technique to generate these multiple looks is to separate the available bandwidth into smaller sub-bands [Moreira, 1990, Scheiber and Bothale, 2002]. These sub-bands are assumed to be equally sized, non-overlapping rect functions in the frequency domain, such that each sub-band is independent and the entire original bandwidth is used. Dividing the original bandwidth B into N_f sub-bands yields a resolution within each sub-band of,

$$\delta_{x,f} = \frac{cN_f}{2B}. \tag{5.2}$$

5.2 ALONG-TRACK RESOLUTION DESIGN IMPLICATIONS

The along-track resolution of a strip-map synthetic aperture sonar system is determined by the number of along-track, or azimuth samples, combined in the synthetic aperture processing algorithm [Hawkins, 1996]. As more along-track samples are combined, the resolution is improved. The coherent summation along the synthetic aperture can only be performed within the extent of the composite along-track beam-pattern of both the transmitter and receiver elements of the sonar. The composite beam-pattern will be

dominated by the larger of the transmitter and receiver elements[1], giving a theoretical along-track resolution, δ_y, of

$$\delta_y = \frac{\max\left[D_t, D_r\right]}{2}, \tag{5.3}$$

where D_t and D_r are the along-track dimensions of the transmitter and hydrophone respectively [Hawkins, 1996]. Defining,

$$D = \max\left[D_t, D_r\right], \tag{5.4}$$

gives the classical theoretical along-track SAS resolution of

$$\delta_y = \frac{D}{2}. \tag{5.5}$$

Like the across-track resolution, this theoretical limit is never achieved from a real-world system, requiring perfect sampling with almost nil platform positional errors for the entire length of the synthetic aperture. Practically, some allowance must be made for resolution loss from a real-world SAS system.

Multiple looks can also be generated for the height estimation by considering along-track neighbouring resolution cells from the reconstructed image. By assuming there is only a small change in the scene for neighbouring reconstructed pixels, these may be considered independent looks of the scene. This does, however, lower the resolution of the resulting image by a factor of the number of pixels combined, giving an along-track resolution of

$$\delta_{y,p} = \frac{DN_y}{2}, \tag{5.6}$$

where N_y is the number of pixels combined.

5.2.1 Along-track sampling requirements

The along-track sampling requirement of a SAS system refers to the along-track distance traveled by each hydrophone between pings. For a single along-track receiver system, this requirement is equivalent to the distance traveled by the entire platform. The along-track sample spacing is normally expressed as a fraction of the largest along-track element size, D. For example, consider a system with a transmitter with an along-track dimension of 0.2 m, and a single along-track receiver with an along-track dimension of 0.1 m. For a forward velocity giving a movement of 0.1 m between pings, the along-track sampling is referred to as $D/2$. With a single hydrophone system (in the along-track direction), the sample spacing can be controlled by altering the pulse repetition frequency (PRF) of the sonar and/or the forward velocity of the platform. For a multiple receiver (in the along-track direction) array based SAS system, the normal mode of operation is for the platform to move forward a distance equal to half the

[1] Assuming the elements are uniform sensitivity and rectangular in shape.

total array length between pings[2]. The resulting receiver positions are therefore fixed in separation, and cannot be controlled by simply altering the speed of the platform. This also implies the along-track sampling ratio is also fixed, determined solely by the size of the transmitter and receiver elements. The along-track sampling ratio for a multiple along-track receiver SAS system is therefore determined solely by the transmitter and receiver element sizes. To achieve the minimum $D/2$ [Tomiyasu, 1978] sampling required, the transmitter size must be of at least twice that of the receiver.

The traditional minimum required sampling ratio of $D/2$ [Tomiyasu, 1978] has been shown to be barely adequate [Curlander and McDonough, 1991, Rolt and Schmidt, 1992], with large grating lobe aliases present in the final image [Callow, 2003]. The level of these grating lobes can be determined from the along-track ambiguity to signal ratio (AASR), defined in Section 3.7.2. For a system with equal along-track sized transmitter and receiver elements, the AASR=-13 dB. The implication is a severe reduction in the processing AASR coherence component to $\gamma_{p(a)} = 0.9501$, calculated from (3.72). A more suitable along-track sampling ratio minimum has been shown to be $D/3$ [Callow, 2003, Hawkins, 1996]. With equal sized transmitter and receiver elements, $D/3$ sampling gives an AASR of -24 dB, yielding a much improved processing AASR coherence component $\gamma_{p(a)} = 0.9952$.

For an interferometric system using an evenly spaced along-track array of receivers the normal mode of operation produces $D/2$ sampling. More dense sampling can be achieved with a multiple receiver array by reducing the along-track speed of the platform. However, only sample spacings of $D/2n$ can be achieved, where n is from the set of integers. Such schemes with sample spacings of other than $D/2$ implies algorithms requiring overlapping phase-centers cannot be used without the use of an unevenly spaced hydrophone array.

KiwiSAS Example

Presented here is a demonstration of the importance of considering the AASR ratio when designing an InSAS system. The dimensions of the system under consideration are based in the KiwiSAS-IV system, parameters of which are shown in Table 2.1. The primary parameters of importance here are the along-track dimensions of the transmitter and receiver elements,

$$D_t = 336\,\text{mm}, \tag{5.7}$$

$$D_r = 232\,\text{mm}. \tag{5.8}$$

[2]Further considerations may also be needed here for redundant phase centres by overlapping of the phase centres of subsequent pings, a requirement for many motion compensations algorithms [Bellettini and Pinto, 2002].

When considering the AASR of the system, the ratio of transmitter to receiver size is important, since this ratio defines the composite transmit/receive beam-pattern of the system. This ratio is dubbed the transmitter receiver along-track ratio, TRAR,

$$\text{TRAR} = \frac{D_t}{D_r}. \tag{5.9}$$

From Figure 3.11 of Section 3.7.2, it can be seen for the equally sized transmitter/receiver case (TRAR = 1) the AASR is -13 dB with D/2 sampling. For the KiwiSAS system, TRAR $\approx$ 1.5, providing a substantial improvement in AASR to -21 dB for D/2 sampling. Similar improvements from and increased TRAR can also be seen in Figure 3.11 for sampling rates of higher than D/2, although the improvement is often less than for the D/2 case. The mechanism for improvement can be seen graphically in Figure 5.1, where the along-track composite tx/rx beam-pattern is shown for the KiwiSAS system specifications, with along-track sample spacing in the range D/2–D/5. As the sampling spacing is increased, the energy folded back into the passband can be seen to decrease. The resulting AASR is plotted in Figure 5.2(a) for various along-track sampling spacings. The KiwiSAS has a typical forward velocity in the range 1.5–3 kn giving sampling ratios in the range of $\approx$ D/4.5–D/2.2. Across this speed range, the AASR shows a small improvement, especially once the sampling is at a rate greater than D/4. The importance of improving the AASR can be seen in Figure 5.2(b), where the AASR is converted to an equivalent coherence component using (3.72). The coherence component can be seen to approach unity for samplings of spacing less than $D/3.5$. The coherence component is altered markedly by small improvements in the AASR, showing its importance when designing an InSAS system.

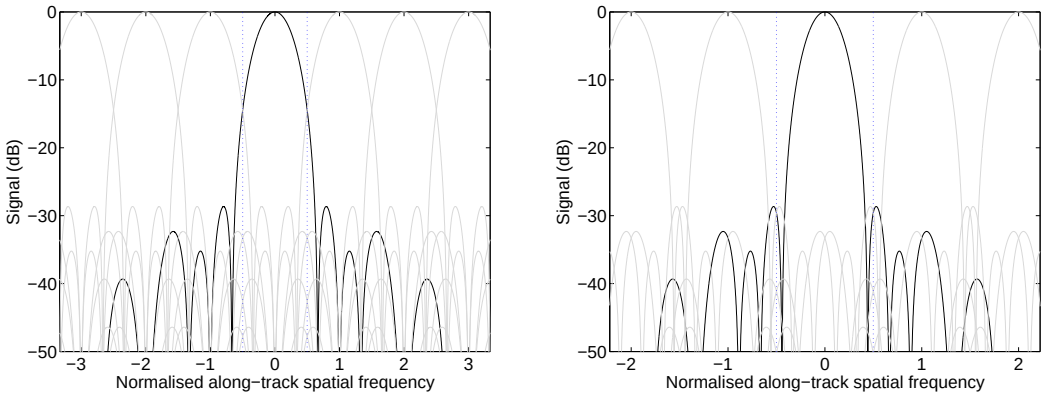
5.3 HEIGHT PRECISION DESIGN IMPLICATIONS

Unlike the across and along-track resolution, the height precision of an InSAS system changes with distance from the imaging platform. This is because for a given error in the time difference as estimated at the receiver array, the error in the estimated height increases with range. For the calculations here a single range is considered, able to be set to any point in the swath as the specifications dictate, shown in Figure 5.3. Assuming the point of interest is in the far-field of the array (incoming wavefronts are parallel across the receiver array), the geometry of Figure 5.4 can be used. Note, here it is assumed the effects of the footprint shift have been corrected for perfectly, allowing it to be neglected here. The difference in path lengths, Δr, can be found using similar triangles of

$$\sin\phi = \frac{\Delta r}{B} = \frac{H}{r}, \tag{5.10}$$

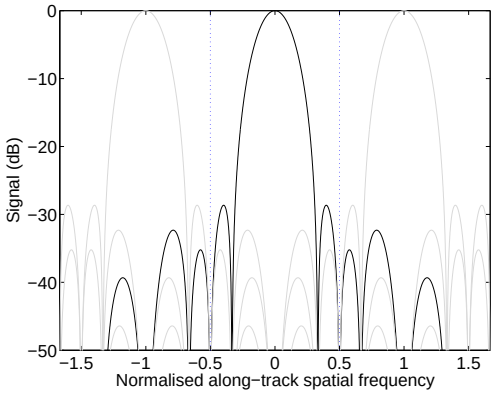
giving a difference in length of

$$\Delta r = \frac{HB}{r}. \tag{5.11}$$

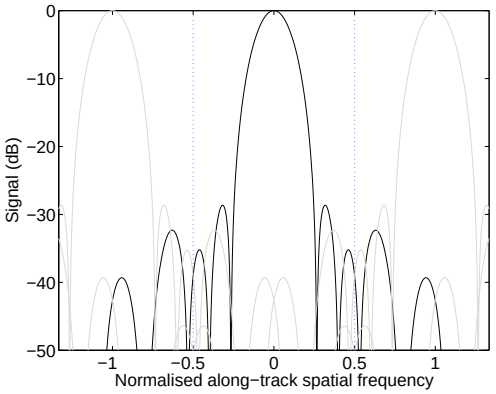


(a) $D/2$.

(b) $D/3$.

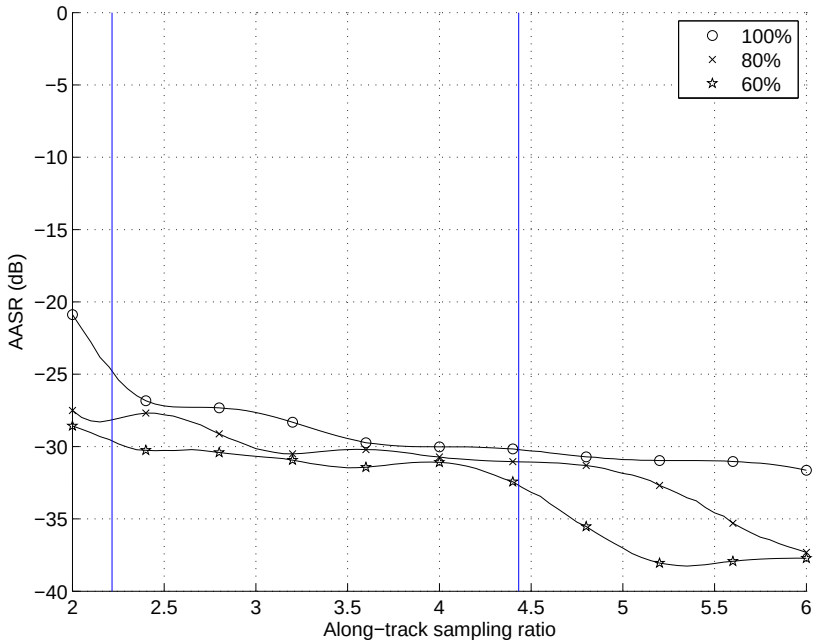


(c) $D/4$.

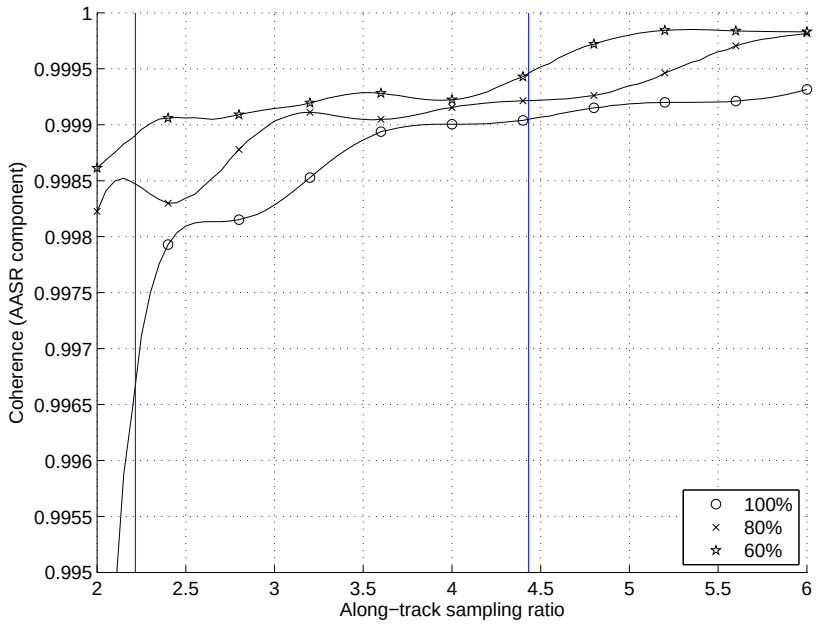


(d) $D/5$.

Figure 5.1 The composite transmit and receive beam-pattern for the KiwiSAS system; transmitter length $D_t = 336\,\text{mm}$, receiver length $D_r = 232\,\text{mm}$. Along-track sample spacing is based on the larger of the two element sizes (transmitter in this case), plotted for spacings of (a) D/2, (b) D/3, (c) D/4, and (d) D/5. The vertical dotted lines at ± 0.5 normalised spatial frequency indicate the full (100%) along-track spatial frequency bandwidth. These plots can be compared to Figure 3.9 where the transmitter and receiver are of equal size.



(a) AASR.



(b) Coherence component $\gamma_{p(a)}$.

Figure 5.2 Based on the parameters of the KiwiSAS system, transmitter length D_t =336 mm, receiver length D_r =232 mm; (a) The along-track ambiguity signal to noise ratio (AASR), and (b) the equivalent coherence component. The region between the vertical solid lines represents the typical operating speed (1.5–3 kn) with a pulse repetition frequency of 15 Hz. When the platform velocity is towards the faster end of the operating region (sampling ratio < D/3) the coherence between the interferometric pairs of receivers can be improved by processing a smaller along-track synthetic aperture. For a sample spacing of D/2.5, processing 60% of the full aperture gives an improved AASR from -27.2 dB to -30.3 dB, increasing the coherence from 0.9981 to 0.9991. By reducing the along-track aperture length the resolution will be degraded, but height estimates will be improved because of the higher coherence. For higher along-track sampling rates (sampling ratio > 3.5), reduction of the along-track aperture length has little effect on the AASR (and hence coherence), so the full aperture should be used to maintain full image resolution. These plots can be compared to Figure 3.10 where the transmitter and receiver are of equal size.

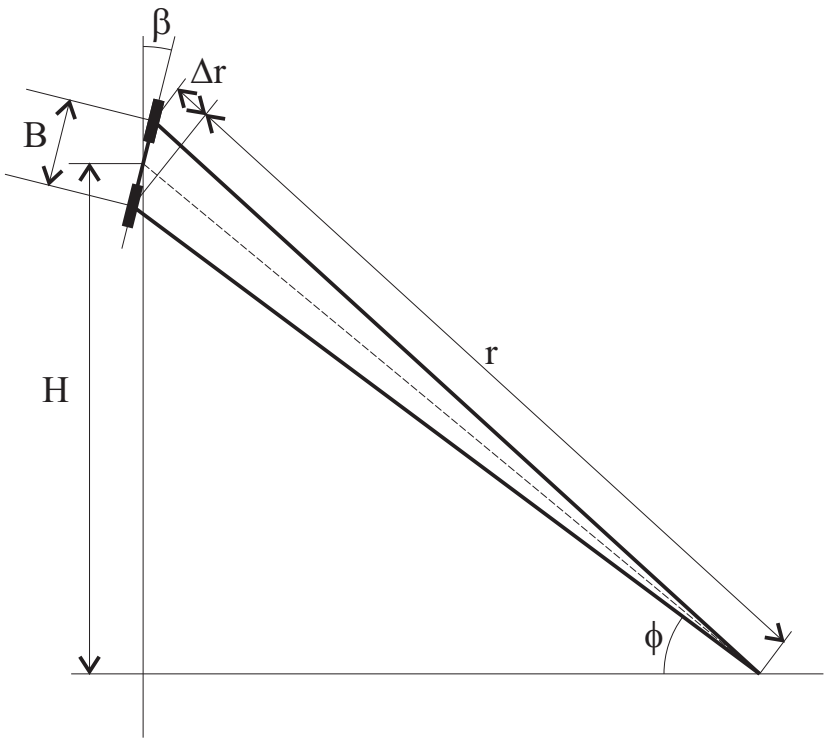


Figure 5.3 Geometry of InSAS sonar system for a single point on the seafloor imaged with two receivers of separation B. Note this is a first order model of the InSAS geometry, neglecting footprint shift effects.

This also assumes the receiver array is in a vertical position, $\beta = 0$, as shown in Figure 5.4. The sensitivity required in the time difference estimate for a given height resolution can be found by considering an object position with a small change in height H, i.e., $H + \Delta h$, giving a range difference of

$$\Delta r + \rho = \frac{(H + \Delta h)\, B}{r}, \tag{5.12}$$

where ρ is the change in range difference due to the small change in object height Δh. After removing the original range difference of (5.11) the accuracy required in the range difference estimation is

$$\rho = \frac{\Delta h B}{r}. \tag{5.13}$$

Alternatively, the height error for a given time difference estimate error is

$$\Delta h = \frac{\rho r}{B}. \tag{5.14}$$

Equation (5.14) implies the height error is directly proportional to the range of the point of interest, and inversely proportional to the baseline length. Both of these parameters are often fixed in a real system, since there is a desired maximum operating range, and the baseline is limited by the size of the sonar vehicle. Thus, a lower limit can be

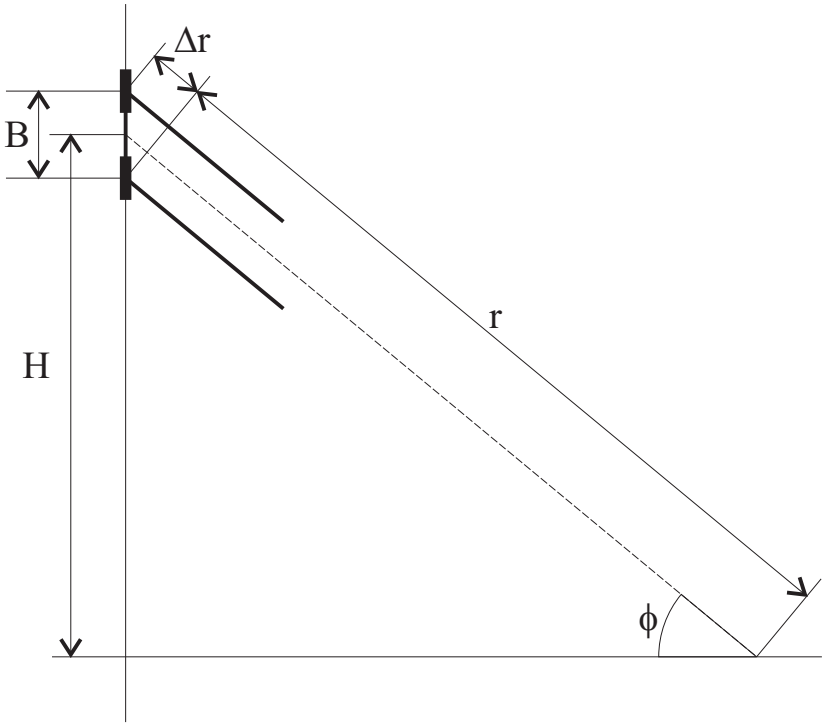


Figure 5.4 First order approximation of the geometry of an InSAS sonar system with a vertical receiver array and field approximations. The far-field approximation assumes the incoming wavefronts are straight such that their propagation paths are parallel, valid for typical side-scan InSAS systems.

calculated for the expected height accuracy, after estimating the maximum expected time difference estimation error. However, the time difference estimation error, ρ, is not independent of r or B since the coherence of the system is dependent on both r and B, a factor in calculating ρ. The height resolution of the system can only be found once the coherence of the system is calculated, and the number of effective looks calculated. Once these are found, Figure 4.5 can be used to find the error in the time delay estimation, and finally the expected height error found from Equation 5.14.

Since the time difference estimation process for a InSAS system is a stochastic process, the height estimation is also a stochastic process. Therefore, the height estimation error of (5.14) will only be satisfied with a given level of confidence rather than a definitive limit.

5.3.1 Coherence and independent looks considerations

To obtain a certain time difference accuracy between the receivers, there must be sufficient coherence between the receivers of the array. This relationship is shown in Figure 4.5 for varying number of independent looks. For a more detailed analysis of interferometric coherence see Chapter 3. In Chapter 4 is was shown multiple estimates of the interferometric phase difference can be combined in a maximum likelihood sense,

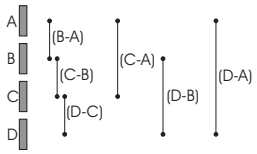


Figure 5.5 Interferometric pairs for a four element vertical array, labelled A-D. A total of six interferometric pairs exist as shown by the vertical bars, however these pairs are not fully independent since they share receiver data values.

yielding one overall estimate. When compared to the true interferometric phase, the overall maximum likelihood estimate will have a lower error variance than any of the individual estimates. Thus, to produce a height map of the scene with small errors, multiple estimates, or 'looks', are required. These looks can be generated in several ways, and fall into two categories; those that do not lower the across- and/or along-track resolution of the image, and those that do lower either or both the across and along-track resolution. The first category is obviously more desirable, however, normally extra hardware is required within the towfish platform.

Looks without lowering image resolution

There are two practical methods to obtain extra looks without lowering image resolution; either by using multiple interferometric pairs, or simultaneously imaging in multiple frequency bands. A third method is to use a multiple pass system, but there is currently no known research in this field. Multiple interferometric pairs can be created from a vertical array of receivers with greater than two elements. For a vertical array of N_h receivers,

$$N_i = \frac{N_h\,(N_h - 1)}{2}, \tag{5.15}$$

interferometric pairs can be derived. However, the number of independent looks from this array does not increase linearly with the number of interferometric pairs. This is because many of the interferometric pairs share the same information from each receiver element and are therefore not fully independent. To find the number of independent looks of the total set of Equation 5.15, the interferometric pairs can be expressed as rows of a matrix. This matrix can then be row reduced to find the minimum independent set. For the example of Figure 5.5, the interferometric pairs of the $N_h = 4$ receivers

can be expressed in matrix from,

$$\begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ -1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix}. \tag{5.16}$$

Via row reduction the coefficient matrix can be reduced,

$$\equiv \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ -1 & 0 & 0 & 1 \end{bmatrix} \equiv \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & -1 & 0 & 1 \end{bmatrix} \equiv \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{5.17}$$

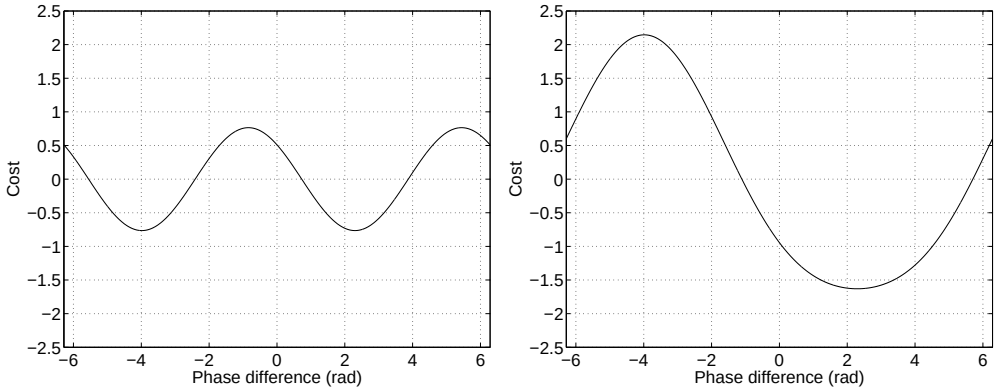
yielding three independent interferometric pairs. More generally, for an array with N_h receivers, the number of independent interferometric pair looks, N_a, is

$$N_a = N_h - 1. \tag{5.18}$$

This is, however, only valid in cases where the phase across the receiver array is linear and continuous. In the case of complex data, this condition may not be satisfied due to the modulo-2π nature of complex data. In typical InSAS setups, this condition can be violated since the maximum baseline is normally several wavelengths, producing 2π phase wraps for certain geometries. This problem can be overcome by using all the interferometric pairs in the time difference estimation, increasing the unambiguous range of the estimation. The use of multiple interferometric pairs can increase the unambiguous estimation range by utilising baselines of different lengths (see Figure 5.6). Since the ratio of the individual baselines is known, the expected time difference across those baselines will also exhibit the same ratio. For more analysis of this see Section 4.6.

Looks by lowering image resolution

By sacrificing resolution in either the across-track or along-track directions, extra looks can be produced to improve the height accuracy. When choosing the resolution loss factor in the across- and along-track directions, several considerations must be made. Since one estimate will be made from a larger portion of the image, the height will be an average estimate over this small sub-portion of seafloor. Secondly, the ratio of



(a) 2 receivers (b) 3 receivers

Figure 5.6 Comparison of cost functions of a three element vertical array using, (a) the two small baseline interferometric pairs, and (b) all three interferometric pairs. The outer pair is not independent of the two inner pairs, with minimum cost at equal phase. In (b) the cost function has a repetition period of twice (a), extending the unambiguous phase interval by a factor of two.

dimensions of the sub-portion should be considered. For display to a human operator it is often desirable to have 'square' pixels. Typically, InSAS sonars have higher resolution in the across-track direction, allowing greater resolution loss in the across-track direction while still producing equal resolution in the two directions. For example, with the KiwiSAS sonar parameters the theoretical across-track resolution is 0.0375 m, and a theoretical along-track resolution of 0.1125 m. Thus, the resolution ratio is approximately 3. To obtain square pixels in the final height estimate, a sub-portion size of 9 across-track resolution cells and 3 along-track resolution cells would yield individual height estimates from a roughly square portion of seafloor of 0.3375×0.3375 m. With these resolution loss factors, each sub-portion height estimation can be made from 27 independent looks.

5.4 INSAS DESIGN EXAMPLE

This section outlines the process to estimate the height accuracy expected from an idealised InSAS system. It is not intended for this InSAS sonar system to be of optimal design, merely an example of the process to estimate the expected height accuracy. Many other detrimental factors are not considered here, such as uncorrectable towfish motion, and sea-surface multi-path effects. As an arbitrarily chosen set of parameters for the resolution of the sonar system, the desired resolution is 0.025 m in each of the two ground-plane dimensions. These resolution goals are for the acoustic backscatter intensity, commonly referred to as the intensity image. It is also desired to have a height estimation with comparable (0.025 m) height accuracy for each resolution cell of the seafloor. The sonar is to operate with an imaging centre frequency of 100 kHz,

a maximum imaging range of 50 m, with an altitude of 10 m above the seafloor. The system is limited to a baseline of 0.15 m, using a single interferometric pair of receivers.

To obtain the desired across-track resolution of $\delta_x = 0.025$ m, the bandwidth required from (5.1) is,

$$\begin{aligned} B_s &> \frac{c}{2\delta_x} \\ &> \frac{1500}{2 \times 0.025} \\ &> 30\,\text{kHz}. \end{aligned} \tag{5.19}$$

For the desired along-track resolution of 0.025 m, the maximum along-track size of either the transmitter of receiver elements is (from (5.3)),

$$\max\left[D_t, D_r\right] < 2\delta_y \tag{5.20}$$

$$< 0.05\,\text{m}. \tag{5.21}$$

The required accuracy limit of the time difference estimate can be found using (5.13). For a desired height estimate error of 0.025 m, receiver baseline of 0.15 m, and maximum range of 50 m, the path difference must be estimated with an accuracy of

$$\hat{\rho} = \frac{0.025 \times 0.15}{50} \tag{5.22}$$

$$= \pm 75.0 \times 10^{-6}\,\text{m}. \tag{5.23}$$

For a centre frequency of f_c =100 kHz, this is equivalent to a phase of ± 0.03142 rad. To achieve this level of time delay estimation error, a combination of high coherence and multiple looks of the scene is required. Reproduced here in Figure 5.7 is a plot of the phase accuracy of the interferometric estimation, for varying coherence. Overlaid is the phase accuracy required for the example design. As can be seen from Figure 5.7, the coherence requirement across the array is high, for the single look case a coherence greater than 0.9995 is required. This is an impractical level of coherence to achieve, necessitating the combination of multiple looks. The number of independent looks required can be estimated once the theoretical level of coherence is found.

5.4.1 Maximum expected coherence

This section briefly estimates the expected maximum coherence between the receivers for the idealised InSAS system under consideration here. The overall maximum expected coherence is a combination of the five coherence components of Section 3.2. The coherence components are calculated for a point 50 m across-track from the sonar, with a towfish altitude of 10 m above the seafloor.

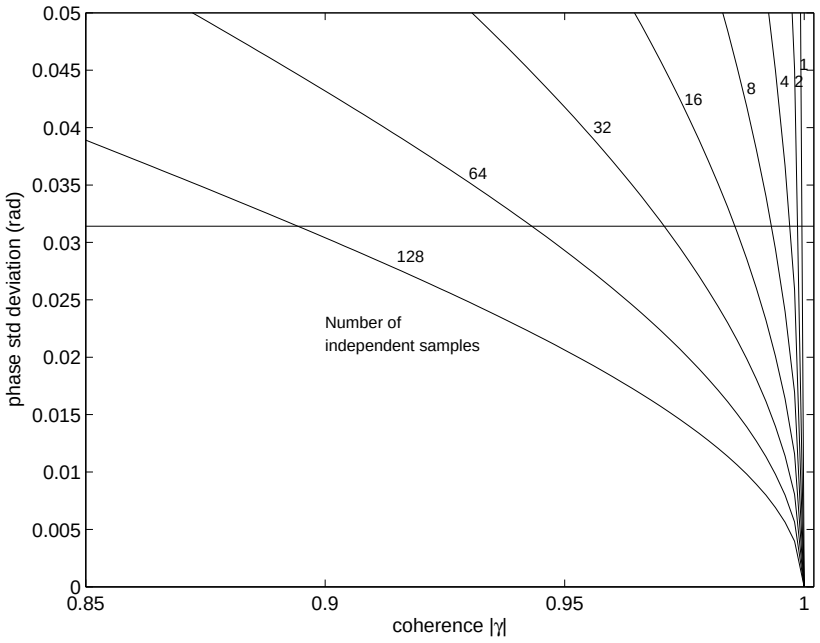


Figure 5.7 Measurement phase difference error variance as function coherence for differing number of looks. Horizontal line represents the phase difference accuracy required for an example 100 kHz centre frequency InSAS sonar with baseline of 0.015 m for an object at 50 m range.

Acoustic noise

In many cases the level of acoustic noise in the system is unknown. However for this section a typical value of $\mathrm{SNR}_n = 40\,\mathrm{dB}$ is chosen, yielding a coherence as shown in Section 3.3 of,

$$\gamma_n = \frac{1}{1 + \mathrm{SNR}_n^{-1}} \tag{5.24}$$

$$= 0.9901. \tag{5.25}$$

Footprint misalignment

Initially, if no correction is made to the SAS data there will be significant footprint shift between the reconstructed images from each receiver. The resulting loss in coherence can be estimated using (3.18), plotted in Figure 3.5,

$$\gamma_m = 0.1774. \tag{5.26}$$

If an accurate prior seafloor height model is known it is possible to re-sample the data from each receiver onto a common surface, eliminating the coherence loss from footprint shifting [Barclay et al., 2003a].

Baseline decorrelation

As the separation of the receivers is increased, there is a greater degree of decorrelation between the measurements. The resulting coherence component γ_b can be calculated as shown in Section 3.5 using (3.49). Assuming the system uses a linear FM chirp, with no spectral shaping, the expected coherence component from baseline decorrelation is given by (3.56),

$$\gamma_b = 1 - \frac{0.15}{51} \tag{5.27}$$

$$= 0.9970, \tag{5.28}$$

where $B_{\text{crit}} = 51$. The baseline decorrelation coefficient will decrease if sub-banding is applied [Lurton, 2000]

Temporal decorrelation

Temporal decorrelation effects will occur if the imaged scene undergoes movement during the acquisition of the interferometric data. Assuming the system is a single pass system, there will be no temporal difference between the receivers. However, movements of the scene could occur during the sonar traversal along the synthetic aperture path. Such scene movements would be rare for a real-world scan of seafloor topography allowing

$$\gamma_t = 1. \tag{5.29}$$

Processing noise

Here it is assumed the data collection is performed perfectly requiring no motion compensation of the data. Therefore all processing noise is considered to be from the standard SAS processing algorithms, namely pulse-compression and azimuth-compression. Standard matched filtering using a linear frequency modulated (LFM) chirp gives a peak to peak side-lobe ratio, SNR_m, of the product of the bandwidth and the transmission time of the chirp. For the signal bandwidth of 30 kHz calculated in (5.20), and a signal transmission time of 10 ms, $\text{SNR}_m = 300$ (49.5 dB),

$$\gamma_{p(m)} = \frac{1}{1 + \text{SNR}_m^{-1}} \tag{5.30}$$

$$= 0.9967. \tag{5.31}$$

Assuming the system operates with a forward velocity resulting in $D/4$ azimuth sample spacing, the AASR is 21 dB [Callow, 2003], giving a coherence component,

$$\gamma_{p(a)} = \frac{1}{1 + \text{AASR}^{-1}} \tag{5.32}$$

$$= 0.9921. \tag{5.33}$$

Combining (5.31) and (5.33) gives an overall coherence component from processing noise of,

$$\gamma_p = \gamma_{p(m)}\gamma_{p(a)} \tag{5.34}$$

$$= 0.9888. \tag{5.35}$$

Overall expected coherence

Combining these effects yields an overall estimate of the expected coherence,

$$\gamma = \gamma_n\gamma_m\gamma_b\gamma_t\gamma_p \tag{5.36}$$

$$= 0.9901 \times 0.1774 \times 0.9970 \times 1.0 \times 0.9888 \tag{5.37}$$

$$= 0.1732. \tag{5.38}$$

This is the maximum coherence expected for the system as specified, without any footprint shift correction. Since in real situations the topography of the seafloor is not known, footprint shift correction cannot be applied initially so Equation 5.38 is valid. Referring back to Figure 5.7 it can be seen this level of coherence, $\gamma = 0.1732$, will require an impractically large number of independent looks (many thousand) to achieve the desired height accuracy. To generate this number of independent looks will require compromised resolution in the across and along-track directions. The most dominant source of coherence loss is due to footprint misalignment, as expected. After corrective data processing, it is possible to increase this coherence component substantially, although this correction does require accurate prior knowledge of the seafloor position. Assuming the seafloor surface is accurately known, the footprint misalignment coherence component can theoretically be increased to 1.0, although more coherence loss will be introduced to the processing coherence component due to interpolation. Assuming this corrective processing can be performed perfectly using a 8-point truncated sinc interpolator (see Section 3.7.3), the overall coherence becomes

$$\gamma = \gamma_n\gamma_m\gamma_b\gamma_t\gamma_p \tag{5.39}$$

$$= 0.9901 \times 1.0 \times 0.9970 \times 1.0 \times (0.9888 \times 0.9990) \tag{5.40}$$

$$= 0.9751, \tag{5.41}$$

where the extra processing coherence component loss of 0.9990 is due to the interpolation. Using the overall coherence of (5.41), and referring again to Figure 5.7, approximately 35 independent looks will be required to obtain the desired height accuracy. To generate this number of full resolution looks, a combination of multiple vertical receivers and multiple frequency bands are required. However, generating in excess of around 10 independent looks requires the use of large vertical arrays and large signal bandwidths, drastically increasing the complexity of the sonar system and data acquisition/storage sub-systems.

As stated previously, correction of the data to the true ground-plane required accurate prior knowledge of the seafloor shape, unknown at this stage of data processing. However, correction to a ground-plane surface closer to the true ground-plane will substantially decrease the mis-registration between the receiver channels, decreasing the footprint misalignment coherence loss component. This prior knowledge of the approximate depth of the seafloor is often known during imaging runs, estimated using some external means. Assuming this initial correction to a nominal seafloor depth is 1 m in error to the true sea-floor depth, the coherence loss due to footprint mis-registration is substantially improved to 0.7960. The overall coherence is now

$$\begin{aligned} \gamma &= \gamma_n \gamma_m \gamma_b \gamma_t \gamma_p && (5.42)\\ &= 0.9901 \times 0.7960 \times 0.9970 \times 1.0 \times 0.9888 && (5.43)\\ &= 0.7770. && (5.44) \end{aligned}$$

Even after correction to a ground-plane close to the true ground-plane, the overall coherence is still poor, requiring in excess of 150 independent looks to obtain the desired height accuracy.

A more practical solution to generate large quantities of looks is by sacrificing resolution in either/both the across- and along-track directions. By reducing both the across- and along-track resolution by factors of 5, 25 looks can be generated for the height estimation process. The effect of resolution reduction, however, has implications on the expected coherence between the receivers. Primarily, reduction of across-track resolution causes a substantial improvement in the footprint misalignment coherence component. Conversely, coherence loss from baseline decorrelation is increased, due to the larger resolution cell size. Using a resolution loss factor of 5, and correction to a ground-plane with height error of 1 m, the expected coherence is

$$\begin{aligned} \gamma &= \gamma_n \gamma_m \gamma_b \gamma_t \gamma_p && (5.45)\\ &= 0.9901 \times 0.9550 \times 0.9927 \times 1.0 \times 0.9888 && (5.46)\\ &= 0.9281. && (5.47) \end{aligned}$$

Referring to Figure 5.7 shows a coherence level of 0.9281 requires approximately 80

independent looks to achieve the desired height accuracy. Although this initially seems to be a large quantity of looks required, 25 are provided by the resolution loss factor of 5 in each direction. Use of three interferometric receiver elements will produce a looks gain factor of 2; a further gain factor of 2 can be obtained using 2 independent imaging frequency bands, such as the configuration of the KiwiSAS system. With a three receiver, dual frequency system, a gain factor of 2×2 is obtained, yielding approximately 100 independent looks of the scene for this example. If extra receivers and/or imaging frequency bands cannot be incorporated into the sonar design, the original aim of 0.025 m height accuracy cannot be achieved without a massive reduction in across- and along-track resolution.

5.5 TOWFISH MOTION

With almost all synthetic aperture sonar systems, the imaging platform undergoes unwanted motion during data acquisition. A total of six parameters are needed to fully describe the position and orientation of an object within a three dimensional coordinate frame. These six parameters are three translational parameters, and three rotational parameters, depicted in Figure 5.8. Each of the three translational motions are parallel to each of the three axes, namely sway, surge and heave. The three rotational motions are pitch, roll, and yaw, defined as rotations around each of the three axes.

5.5.1 Motion measurement

The ideal tow-fish navigation system would provide all six position/motion parameters at every ping position along the synthetic aperture. However, due to physical placement within the towfish body, individual measurements are not independent of each other. For example, consider the case of towfish equipped with a perfect set of sensors capable of perfectly estimating all six parameters, while experiencing a pitch in the positive direction around the central point of the towfish (see Figure 5.8). Assuming all the sensors are located exactly at the centre of the towfish, the pitch will be measured correctly by the pitch sensor. However, if all the sensors are located behind the centre of rotation, the pitch will still be measured correctly, but the heave sensor will also measure a negative translation. Also to a lesser degree, the surge sensor will measure a small positive translation. The measurement of motion is further complicated when the sensors for each of the parameters is located in differing positions within the towfish.

Because of the hydro-dynamic properties of the towfish body, the centre of rotation is difficult to locate, and can be different for each of the three axes. The centre of rotation may also vary depending on the mechanism causing the motion. For example, in a towed system, a motion from the tow cable can exhibit different behavior to that of a water column induced movement.

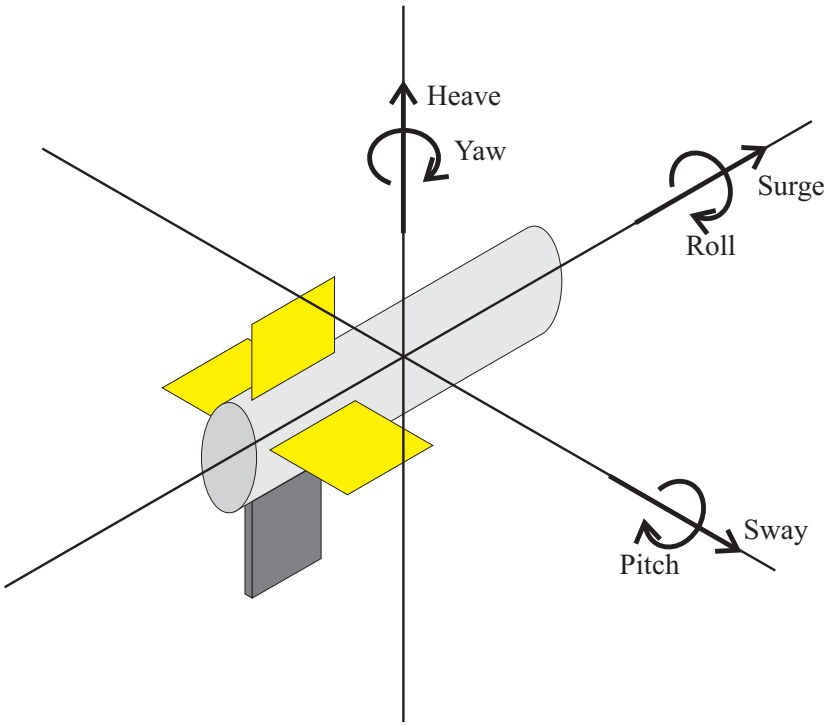


Figure 5.8 Directions of motion. Three translational motions along each of the three axes, and three rotational motions around each of the three axes.

Gravity referenced sensors

One of the simplest devices for measuring pitch and roll is a gravity based sensor. These sensors often consist of a small liquid filled chamber, which under the influence of gravity remains level. The orientation of the chamber can then be found by measuring the level of the liquid within the chamber. The properties of the liquid within the chamber greatly influence the characteristics of the measurements made from the sensor. By using a more viscous liquid, the rate of change of the sensor is decreased, thus limiting the maximum frequency of movement accurately measured. A viscous liquid will, however, exhibit less overshoot when subjected to a fast movement. Since these sensors are gravity referenced, only pitch and roll can be measured.

Accelerometers

Solid state accelerometers have greatly reduced in size and cost, while still maintaining high accuracy. By utilising three sensors, each aligned to the axes directions it is possible to measure accelerations in each of the three directions. By integration over time, velocity in each of the three directions can be obtained. A second integration then yields position. This double integration process, however, can accumulate large errors for small errors in acceleration measurements. With an InSAS system, the data

required for the synthetic aperture can take a significant time, yielding large positional errors from the double integration process.

Further problems with accelerometers are encountered when the towfish motion consists of rotations as well as positional movements. Since the acceleration measurements are recorded along the axes of the towfish, any rotation of the towfish mis-aligns these measurements with the desired global coordinate system. If there is a mis-alignment of the towfish to the global coordinate system, the measured accelerations must be projected back onto the global system.

Magnetometers

Magnetometers can be used to measure the velocity through the Earth's magnetic field. By utilising three sensors, each aligned to the axes of the towfish, the towfish velocity in each of these directions can be measured. The velocity measurements can then be integrated over time to give a position estimate. However, this integration process can introduce cumulative errors, especially when the integration is performed over long periods of time. Also local fluctuations in the earths magnetic field can cause false velocities to be measured as the sensors are moved through the fluctuation. Such fluctuations are common in iron rich soil, and also near undersea power cables, as often found in shallow harbours.

Ring laser gyroscopes

Ring laser gyroscopes consist of a laser beam split into two beams, each of which are reflected in opposite directions around a ring[3]. If the system is rotated around the normal to the ring, the two opposing laser beams experience a Doppler shift. For the beam traveling in the same direction as the rotation the Doppler shift is an increase in frequency, for the opposite direction a decrease in frequency. The difference in frequency for the two laser beams is measured using an interferometer, with the frequency difference proportional to the rotational velocity. Integration over time yields a rotational position, or orientation. Recent advances in ring laser gyroscope technology have greatly reduced the size of the sensor, allowing their use in small towfish bodies.

Doppler velocity log (DVL)

Doppler velocity logs (DVL) are acoustic instruments mounted on the underside of the vehicle, pointed downwards. Their primary use is to provide high accuracy velocity measurements, and altitude from the seafloor. A DVL is particularly suited to autonomous underwater vehicles [Schock et al., 2005, Lee et al., 2004], exhibiting low power consumption and high accuracy measurements. A DVL operates by transmission

[3]Often the ring is a triangular shape, requiring three mirrors.

of a waveform towards to seafloor on an angle from nadir. The reflected waveform from the seafloor is recorded, the Doppler shift giving a measure of the platform speed. By using several angled beams, the platform velocity and heading can be determined.

5.5.2 Roll Correction

The accurate estimation of towfish roll is important for InSAS systems [Bonifant et al., 1999, Bonifant, 1999, Griffiths et al., 1997]. Since the object height is inferred from an angle of arrival of the reflected sound energy, any roll of the towfish body will create an error in the estimated height. Because of the range magnification, even small roll errors will produce large height errors. With a towfish roll β, and an across-track object range of x_0, the first order height error, Δh_{roll} is

$$\Delta h_{\text{roll}} = -x_0 \sin \beta. \tag{5.48}$$

The negative sign is to obey the coordinate system as defined in Figure 5.8, where a positive roll will increase the depth of the height estimate. For example, a 1° roll with an object at 50 m across-track range has an intrinsic height error of -0.873 m.

Rather than estimate the roll using instrumentation, the roll can be estimated directly from the data[4]. If the imaged scene is nominally flat, the roll of the towfish can be estimated by calculating the expected phase difference for each range bin, and comparing with the actual phase difference between the receivers. By then taking a weighted least mean squares estimate of all the range bins, a roll estimate can be made for each ping.

This technique relies on correct estimation of the towfish height about the seafloor, and that the average height across the swath is equal to the height of the towfish. Estimating the height of the towfish could be determined using an acoustic bottom profiler mounted on the bottom of the towfish, or directly from the data [Hayes and Ho, 2000].

Experimental results have shown the data driven approach to roll estimation produced similar results to using a gravity referenced clinometer within the towfish. An example of the estimated roll is shown in Figure 5.9 from a sea-trial with the KiwiSAS sonar system in Lyttelton harbour.

5.6 HEIGHT ESTIMATE PERFORMANCE

The use of a test scene with known height allows the performance of a height estimation algorithm to be quantified. With a quantifiable measure of performance, comparisons

[4]The method and results presented here was originally published in "Test Results From a Multi-Frequency Bathymetric Synthetic Aperture Sonar", by M. P. Hayes, P. J. Barclay, P. T. Gough, and H. J. Callow, appearing in the Oceans 2001 Proceedings, Vol. 1, Pages:1682-1688, November 2001.

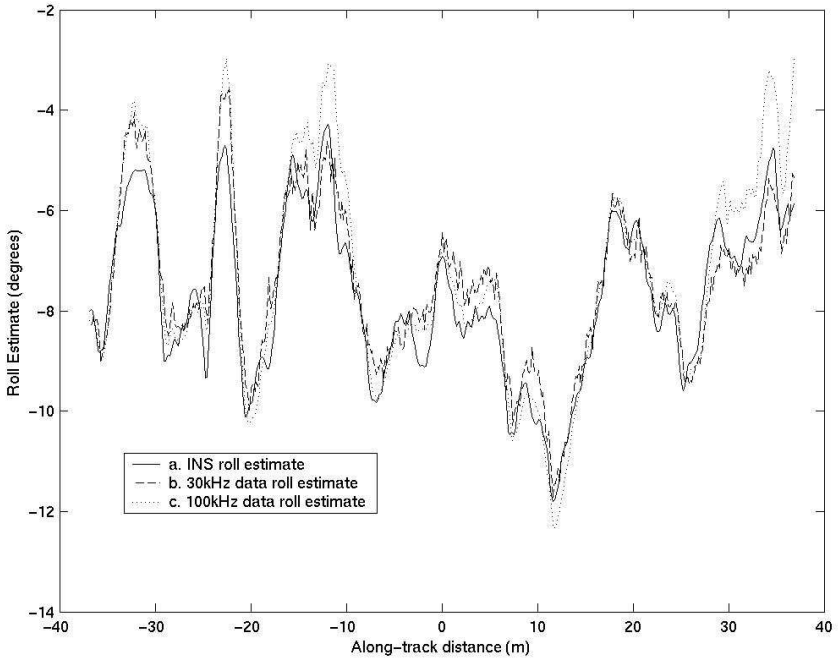


Figure 5.9 Towfish roll estimated from the two frequency bands, 30 kHz and 100 kHz each of 20 kHz bandwidth, from a sea-trial of the KiwiSAS sonar system in Lyttelton Harbour. Also shown is the clinometer roll estimate, showing close agreement with the two data-driven roll estimates.

can be made between algorithms, and also as parameters of the system are changed. In this section various processing parameters are investigated using a test scene based on the KiwiSAS specifications.

5.6.1 Simulator

A major difficulty in developing and testing SAS and InSAS algorithms is the lack of ground-truth information for collected data sets. When data is collected from a sea-trial, the exact specifications of the imaged scene are unknown. Although known man-made objects can be introduced into the test scene, their exact location and orientation is unknown. Also, the surrounding terrain cannot be accurately determined, except for some simple parameters such as average water depth and expected material composition. Another major problem with field collected data is image degradation from unwanted towfish movement. The towfish motion can be controlled through use of a rail based system but these systems are expensive to build and operate, also requiring a suitable location.

Simulation allows for controlled imagery to be generated, with known ground-truth information. Algorithm sensitivity to operational parameters can also be investigated easily, via multiple simulations with one of the many parameters varied. An example is the sensitivity of InSAS algorithms with varying receiver baselines. The effect of variation of the receiver baseline would be extremely difficult to quantify with sea-trials with a large associated hardware cost.

The simulated scenes used in this section are generated using a rough facet Synthetic Aperture Sonar simulator, developed by Alan Hunter of the University of Canterbury[5] [Hunter et al., 2003, Hunter and Hayes, 2005, Hunter, 2006]. This simulator uses a facet based model to describe the imaged scene, allowing complicated shapes to be simulated. Facet shadowing is resolved using ray-tracing techniques, as are multi-path effects from the sea surface. The facet based approach is computationally expensive, so a parallel computing implementation is used by dividing the overall scene into smaller jobs. The simulator is coded in the Python[6] language, with LAM-MPI[7] for parallel computing. Two computing facilities were used for scene simulation, firstly a mosix[8] based cluster of dual processor PCs, and secondly a grid based cluster of dual processor PCs. Both of these clusters are facilities of the Electrical and Computer Engineering Department of the University of Canterbury.

The simulated scene shown in Figure 5.10 is used throughout many of the following sections as test scenes. The scene is cropped into two distinct areas as shown in Figure 5.10, one containing just the flat rough seafloor (dubbed 'background'), the

[5] Electrical and Computer Engineering Department
[6] www.python.org
[7] Local Area Multi-computer - Message Passing Interface, www.lam-mpi.org
[8] www.mosix.org and openmosix.sourceforge.net

second an area containing the aircraft and shadow regions (dubbed 'stuka'). These two scene subsets were chosen since they can be used to illustrate various algorithms performance in areas of complicated scene using the 'stuka' portion compared to the reference 'background' portion.

5.6.2 Height performance metric

The metric used for the following sections is based on the root mean square (RMS) error between the height estimate and the reference. For a reference scene height $\mathbf{h}$, and a scene height estimate $\hat{\mathbf{h}}$ the RMS error is,

$$h_{\text{rms}} = \sqrt{\frac{\sum^{N_x,N_y} \left(\hat{\mathbf{h}} - \mathbf{h}\right)^2}{N_x N_y}}, \tag{5.49}$$

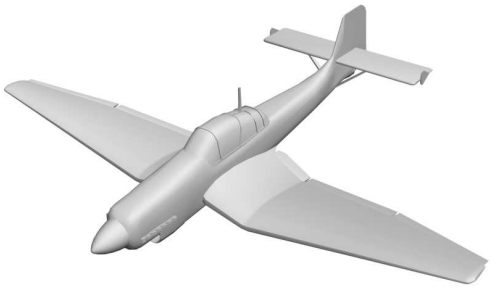
where N_x and N_y are the scene pixel dimensions.

When comparing algorithms or parameter selection it is important to isolate the effect on the height estimate error. This is not always possible due to an interaction within various parameters, but should be minimised and noted when comparisons are made.

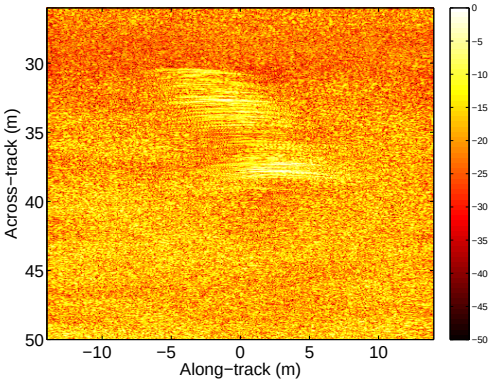
5.6.3 Sub-banding into multiple narrowband signals

An example of sub-banding of SAS imagery is shown in Figure 5.11 where the original bandwidth of 20 kHz is sub-banded into 13 subbands. Distinct differences are seen between the subbands, although each band has equal theoretical across-track resolution. The theoretical across-track resolution of the original broadband imagery as given by (5.1) is 37.5 mm, increased to 488 mm for each of the 13 subbands.

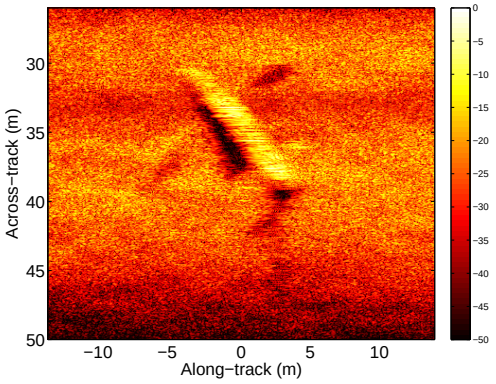
Height estimate errors from different number of subbands are shown in Figure 5.12 for no additive noise, and Figure 5.13 with additive noise. In all cases, the width of the bandwidth of each sub-band is altered such that the entire original spectrum is covered. As expected, as the number of subbands is increased from one, the RMS error rapidly decreases. As the number of subbands is increased there are two mechanisms yielding the reduction in error. The first mechanism is the reduction of bandwidth within each of the subbands, increasing the overlap between the resolution cells of each receiver, decreasing footprint mis-registration (see Section 2.3 and Section 3.4). The second mechanism reducing the error is due to an increased number of independent looks of the scene (see Section 4.3). As the number of looks is increased there is a counter mechanism increasing the error, caused by the loss of across-track resolution. By increasing the resolution cell size with increased looks, the resulting height estimate at a given point will be the average height of the scene over the entire resolution cell. Thus, if the scene contains areas of rapidly changing height the resulting estimated



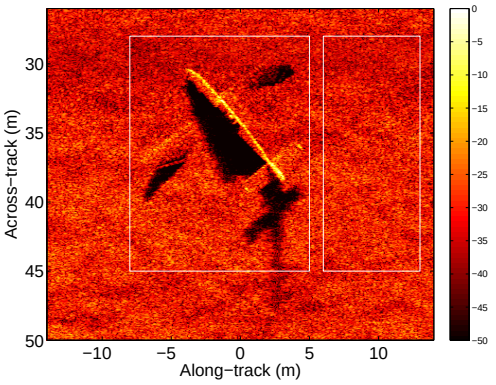
(a) 3D facet based model of 'Stuka' aircraft.



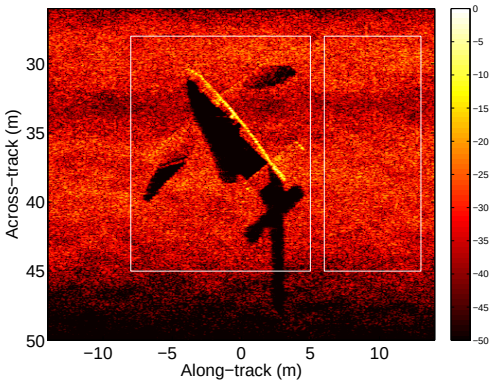
(b) Simulated echo response, f_c=30 kHz.



(c) Simulated echo response, f_c=100 kHz.

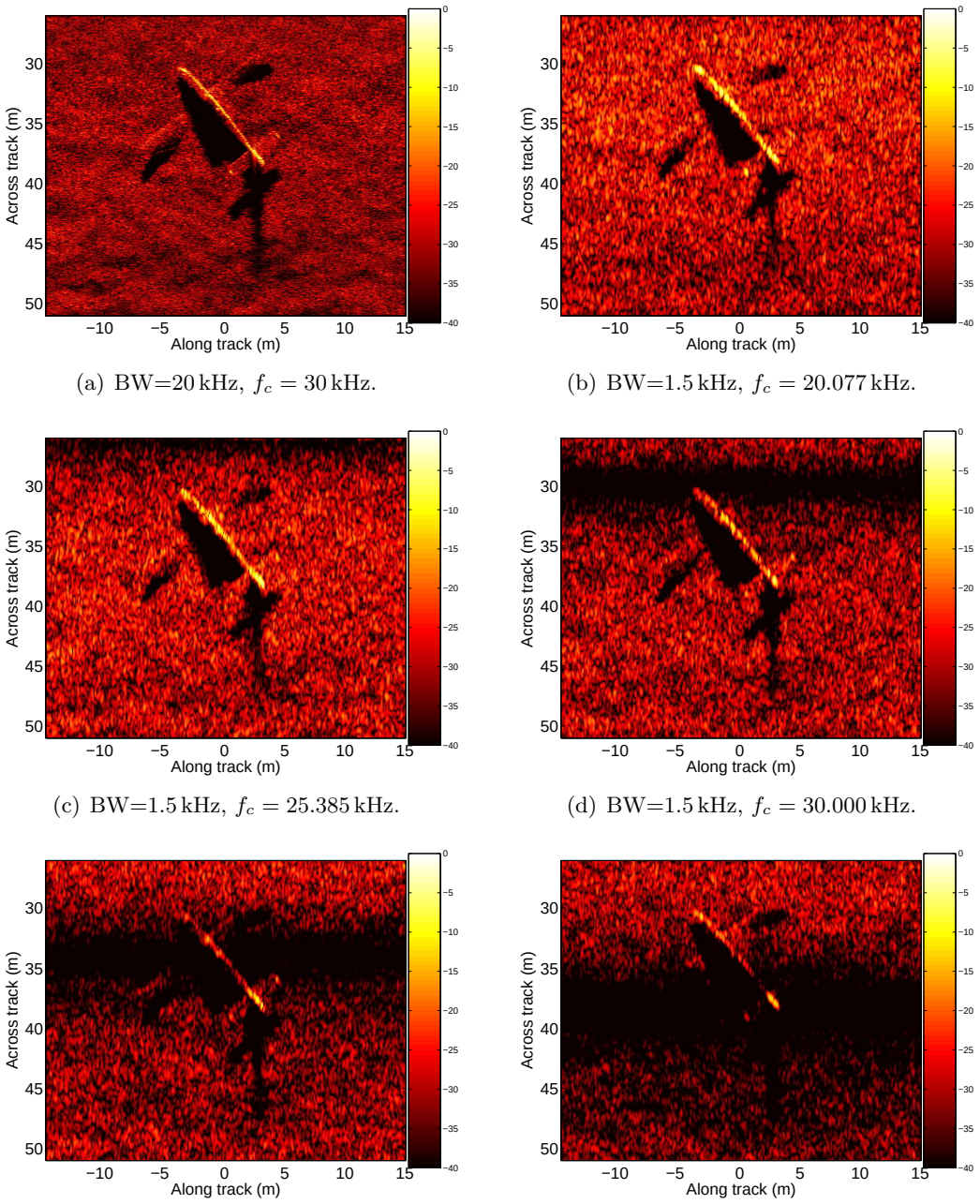


(d) Reconstructed scene, f_c=30 kHz.



(e) Reconstructed scene, f_c=100 kHz.

Figure 5.10 Example of simulated Synthetic Aperture Sonar imagery. The scene consists of a Junkers Ju-87 "Stuka" aircraft partially buried in a nominally flat, undulating seafloor. The sonar is simulated with two simultaneous imaging bands, centred on f_c =30 kHz and f_c =100 kHz, each with a bandwidth of 20 kHz. 5.10(a) shows a 3D view of the stuka aircraft, with the raw simulated imagery for the lower frequency band shown in 5.10(b), and upper band in 5.10(c). After synthetic aperture reconstruction, 5.10(d) and 5.10(e) are produced for the two operating bands. The squares in the reconstructed imagery show the portions of scene dubbed 'background' and 'stuka'.



(a) BW=20 kHz, $f_c = 30$ kHz.

(b) BW=1.5 kHz, $f_c = 20.077$ kHz.

(c) BW=1.5 kHz, $f_c = 25.385$ kHz.

(d) BW=1.5 kHz, $f_c = 30.000$ kHz.

(e) BW=1.5 kHz, $f_c = 34.615$ kHz.

(f) BW=1.5 kHz, $f_c = 39.923$ kHz.

Figure 5.11 The original imaging bandwidth can be divided into several subbands to give multiple lower resolution looks of the scene. 5.11(a) is the normalised intensity image from the original 20 kHz imaging bandwidth, 5.11(b)-5.11(f) show five of the 13 subbands, each with a bandwidth of approximately 1.5 kHz, normalised to a common value across all 13 subbands. The subbands have an across-track resolution cell of 13 times greater size than the original, evident along the plane fuselage where single bright points are smeared out in the across-track direction. The vertical beam-pattern null becomes more significant in the narrowband images giving large low amplitude bands across the imagery.

heightmap will be a smoothed average yielding a large RMS error. This mechanism can be seen in Figure 5.12, where the error is seen to increase for more than approximately 15 looks.

5.6.4 Multiple frequency bands

For InSAS systems with multiple frequency bands, height estimation performed using the highest frequency band should produce the lowest variance in the height estimate for a given coherence. However, since the phase difference between two receivers can only be measured modulo-2π, multiple phase wraps will be present in the measured phase difference. These phase wraps will be more prevalent with higher centre frequencies, as illustrated in Figure 5.14. Two-dimensional phase unwrapping can be used to overcome this problem in some cases but there still remains a height ambiguity in most cases. By combining information from more than one frequency band when estimating the time delay between receivers, the height can be inferred unambiguously with only loose a priori bounds applied to the height. Two possible methods to combine these frequency bands are outlined in the next two sections.

Bootstrapping from the lower frequency band

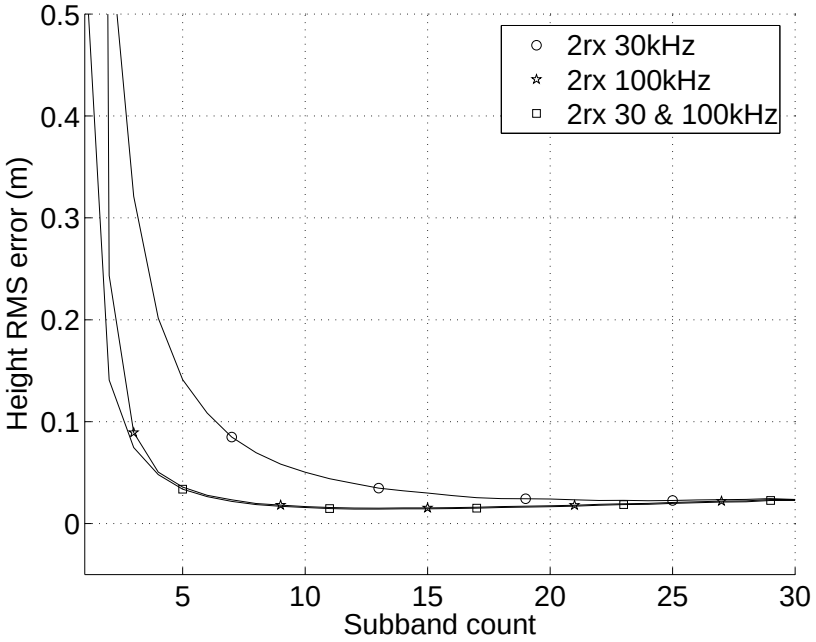
Since the lower frequency band can be used to give an unambiguous height estimate over a wide range of heights, this estimate can be used to bootstrap the height estimate of the higher frequency band. This process can be seen in Fig. 5.14 where the lower frequency band can be used to give an unambiguous height estimate for heights between the 0 th and 1 st zero phase lines. This estimate can then be used to calculate the number of 'unwraps' to apply to each point in the higher frequency estimate. The expression used to find this unwrapping count for a pixel at range r is given by

$$m = \text{round}\left(\frac{2d\left(\hat{z}_1 - \hat{z}_2\right)}{\lambda_2 r}\right), \tag{5.50}$$

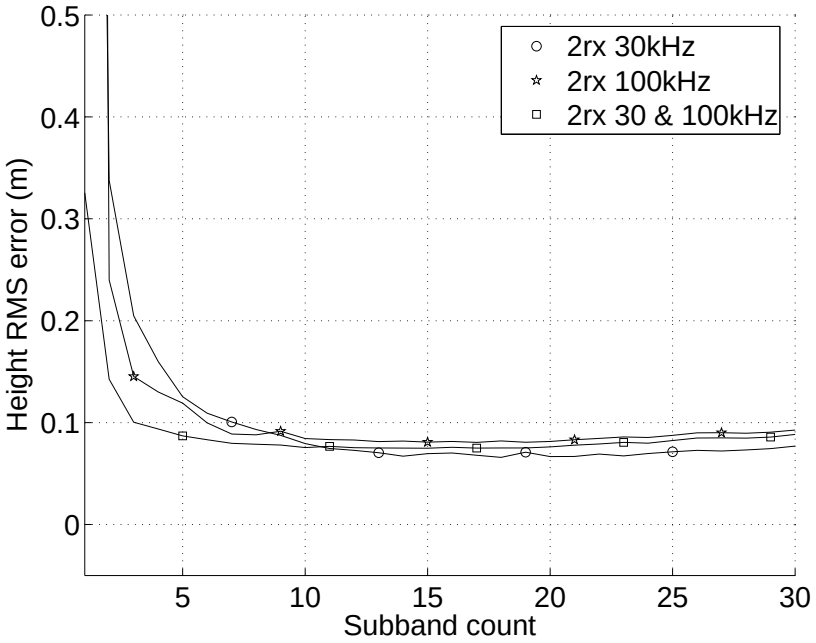
where $\hat{z}_k$ is the height estimate from the k^{th} frequency band and d is the hydrophone spacing. This unwrap count m can then be used to unwrap the higher frequency height estimate to give the bootstrapped height estimate

$$\hat{z}_b = \hat{z}_2 - \frac{m\lambda_2 r}{2d}. \tag{5.51}$$

An example of this process is shown in Figure 5.15. Utilising just the upper frequency band gives a height estimate with high variance, with many pixels reconstructing to an incorrect height. These errors are due to the loose height bounds placed on the height during the height reconstruction, allowing several possible height positions due to the short wavelength. However, by utilising the height of the lower frequency band,

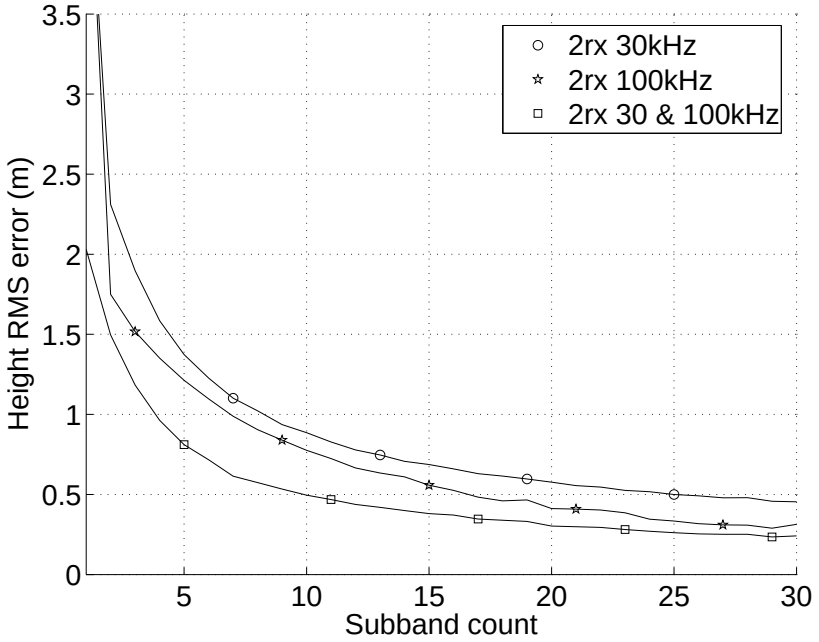


(a) Background portion.

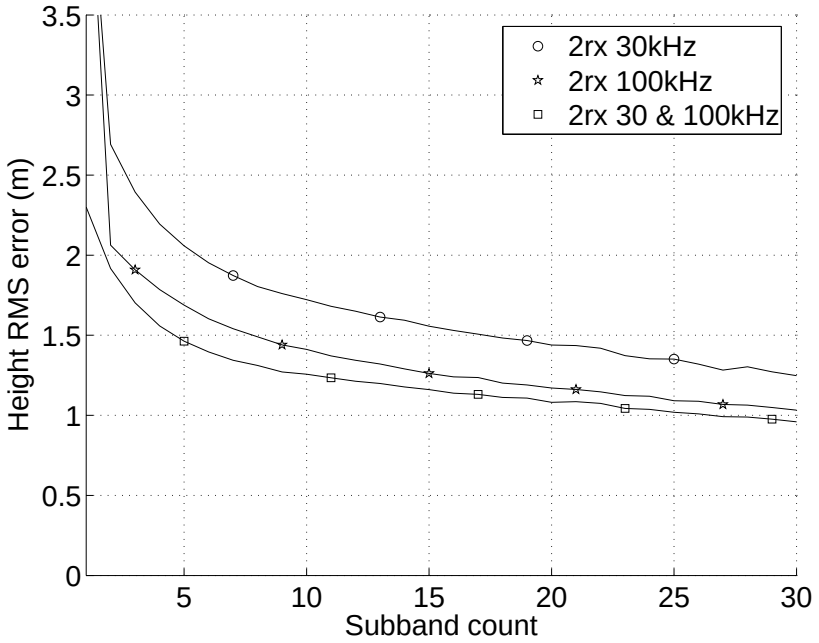


(b) Stuka portion.

Figure 5.12 RMS error of bathymetric height estimates for two portions of the test scene; Figure 5.12(a) for the 'background' portion, Figure 5.12(b) for the 'stuka' portion. No additive noise. As the number of sub-bands are increased the RMS error decreases. Little difference in error is present between the 2 receiver case and 3 receiver case.



(a) Background portion.



(b) Stuka portion.

Figure 5.13 RMS error of bathymetric height estimates for two portions of the test scene; Figure 5.12(a) for the 'background' portion, Figure 5.12(b) for the 'stuka' portion. Noise added with RMS power 20 dB below the peak of the pulse compressed echo response. As the number of sub-bands are increased the RMS error decreases.

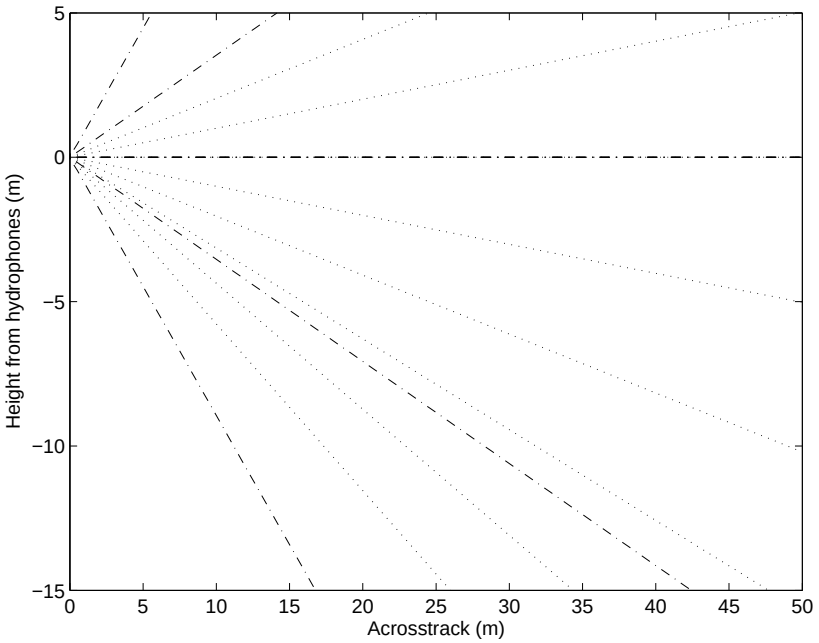
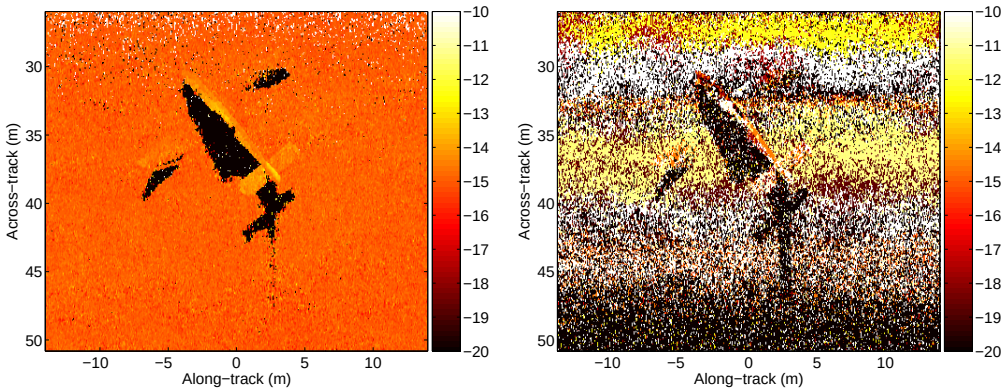


Figure 5.14 Positions within the water column from which the returned echoes will have zero phase difference between receivers. Two centre frequencies are considered; 30 kHz (dash-dot), and 100 kHz (dotted) each with a receiver spacing of $d = 0.15$ m.

the height ambiguity of the higher frequency band can be overcome. Since the lower frequency band is only used to phase unwrap the higher frequency band, the resulting height map variance is a function of the higher frequency data only. If the height estimate from the lower band is incorrect, the high frequency band will be bootstrapped to an incorrect position, giving an error in the final height estimate.

Direct delay estimation from multiple frequency bands

Rather than estimating the height from each frequency band individually and combining the individual height estimates, one height estimate can be made such that it satisfies both frequency bands simultaneously. The direct estimate can be made in a maximum likelihood sense by assuming the frequency bands are independent, allowing the cost functions to be summed as in (4.71). This composite cost function is then searched for the minimum, giving the most likely time difference and hence the most likely height estimate. This direct method is advantageous over the bootstrap method, since the data point from all frequency bands is used in the final height estimate. This method increases the number of interferometric looks of the scene, decreasing the variance of the height estimates. Alternatively, the number of looks from each individual frequency band could be reduced, still giving the same overall number of looks and hence height variance in the final height estimate. A comparison of the direct method and the bootstrap method is shown in Figure 5.15, with superior results in regions of



(a) 30 kHz band.

(b) 100 kHz band.

(c) 100 kHz bootstrapped from 30 kHz.

(d) 30 kHz and 100 kHz direct.

Figure 5.15 Height estimates of the scene can be improved by using multiple frequency bands. Processing the bands individually gives the height estimates of 5.15(a) for the f_c =30 kHz band, and 5.15(b) for the f_c =100 kHz band. Significant height variations can be seen in the f_c =30 kHz band in areas where the surface is smooth, for example the tail section of the aircraft. The f_c =100 kHz height estimate has significant problems due to phase-wrapping ambiguities across the receiver array. Two methods for combining the two frequency bands are shown in 5.15(c) for a bootstrap method and 5.15(d) for a direct multi-frequency estimate.

low signal strength (due to beam-patterning effects) in one of the two frequency bands.

5.6.5 Multiple interferometric pairs

Two separated, simultaneously sampled receivers with overlapping beam-patterns are considered an interferometric pair. For a given time delay, this single receiver pair gives one look of the interferometric phase of the incoming wavefronts. If more than two receivers are simultaneously sampled, multiple looks of the interferometric phase can be obtained from each pair of receivers within the array. However, as shown in Section 5.3.1, each of the looks are not independent of each other. The effective number of looks, N_a, for an array of N_h receivers is given by (5.18), namely

$$N_a = N_h - 1. \tag{5.52}$$

By increasing the number of looks of the scene the height error variance can be reduced, as shown in Figure 4.5. When a low number of looks are used the variance can be large, effectively hiding some objects of interest within the scene.

Unevenly spaced receiver arrays

A problem common to many single receiver pair interferometric systems is producing an absolute phase difference from modulo-2π phase measurements. Whenever the baseline is longer than the wavelength of the incoming wavefronts, this problem can occur. For a measured phase difference $\hat{\phi}$, the absolute phase difference ϕ is one solution of,

$$\phi = \hat{\phi} + n2\pi \qquad n \in (-N_w \cdots N_w)\,, \tag{5.53}$$

where N_w is the lowest integer ratio of baseline length B to wavelength λ,

$$N_w = \text{floor}\left(\frac{B}{\lambda}\right). \tag{5.54}$$

Thus, for a baseline length greater than $\frac{1}{2}\lambda$ there will always be more than one possible solution, only one of which is the true absolute phase difference. Under some circumstances, this ambiguity can be resolved with prior information, allowing limits to be placed on the expected absolute phase difference. This reduces the solution space, leaving only one possible (correct) solution. For most InSAS system geometries this concept can be used to reduce the solution space to those angles below horizontal. This restriction is assuming the imaged scene is always of greater depth than the sonar platform.

The use of an unevenly spaced receiver array gives an increase in unambiguous phase unwrapping range.

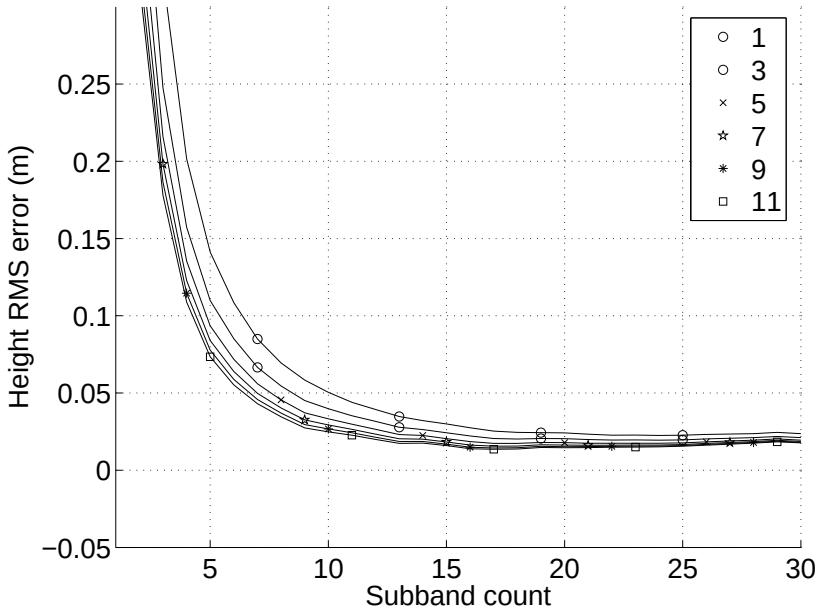
5.6.6 Multiple along-track (azimuth) looks

Combination of multiple along-track looks can be achieved as outlined in Section 4.4.4. As the number of looks is increased the height error variance will decrease. Height error curves for the two scene test portions are shown in Figure 5.16. As expected, the time delay estimation variance initially improves with the use of multiple azimuth looks. However, there is a diminishing gain effect as the number of looks is increased. Initially large improvements are made, with decreasing gains for high numbers of azimuth looks. When additive noise is added to the test scenes, the improvement with greater number of azimuth looks becomes greater as shown in Figure 5.17. Again, a diminishing gains effect can be seen with the RMS error tending toward a limit with increasing look count.

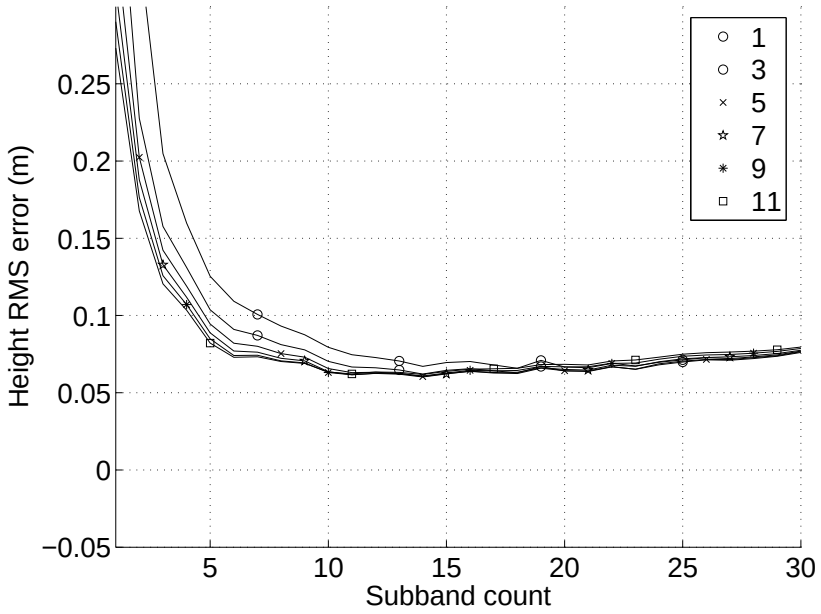
5.6.7 Shadow region detection and masking

In a shallow water environment, sonar imaging systems must operate at a shallow look angle. As a result, any object proud of the seafloor will cast a shadow. The shadow footprint on the seafloor has a significantly different shape depending on the nature of the sonar system, as shown in Figure 5.18. For a real aperture sonar with a narrow along-track beam-pattern, the shadow region has an along-track extent of similar size to the shadowing object, primarily consisting of a fully shadowed umbra. For a synthetic aperture sonar system the along-track beam pattern is wider, and thus the shadow region behind the object covers a larger area. However, the shadow region will primarily be a partially shadowed penumbra region, with a small umbra region directly behind the object.

The consequence of the shadow is to lower the signal strength within the shadow region, and hence lower the SNR. The lower SNR causes a decorrelation across the receiver array, resulting in poorer height estimates. In the fully occluded umbra shadow region all signal is lost, with meaningless resulting height estimates. Ideally this shadow region should be detected and ignored in the final height reconstruction. One method for shadow detection is using the value of the minimum found in the cost function for each pixel. In shadow regions the coherence between the hydrophones is very low, giving a higher cost minimum when compared to non shadow pixels. Pixels falling below the threshold are considered to be in shadow and are then used to mask the height image. The threshold level for optimum shadow masking depends on many factors including; coherence estimate window size, coherence estimator algorithm, signal to noise ratio, etc. For the shadows presented here a shadow mask coherence threshold of approximately 0.75 is used. It should be noted the shadows presented here may appear more defined than those from a real synthetic aperture sonar platform due to motion effects and multipath from the sea surface [Hunter and Hayes, 2005].

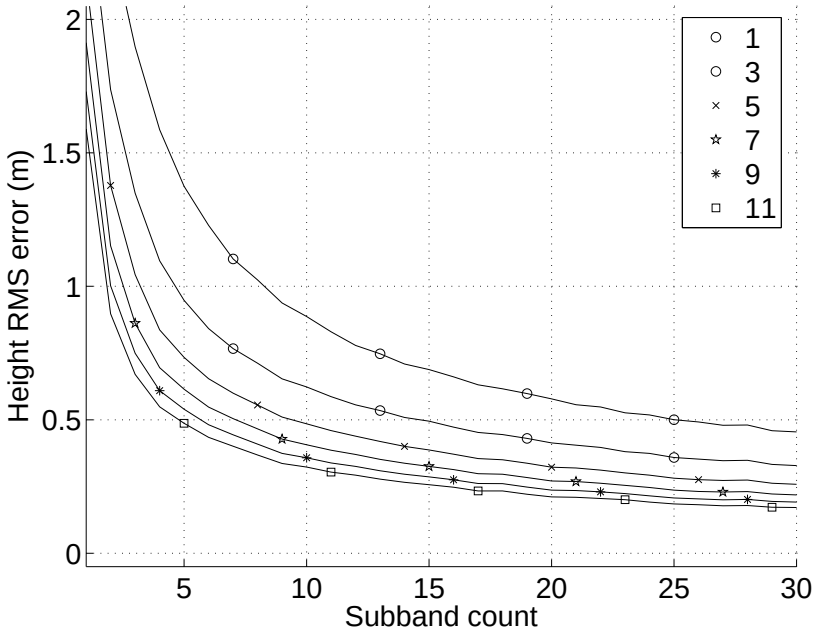


(a) Background portion.

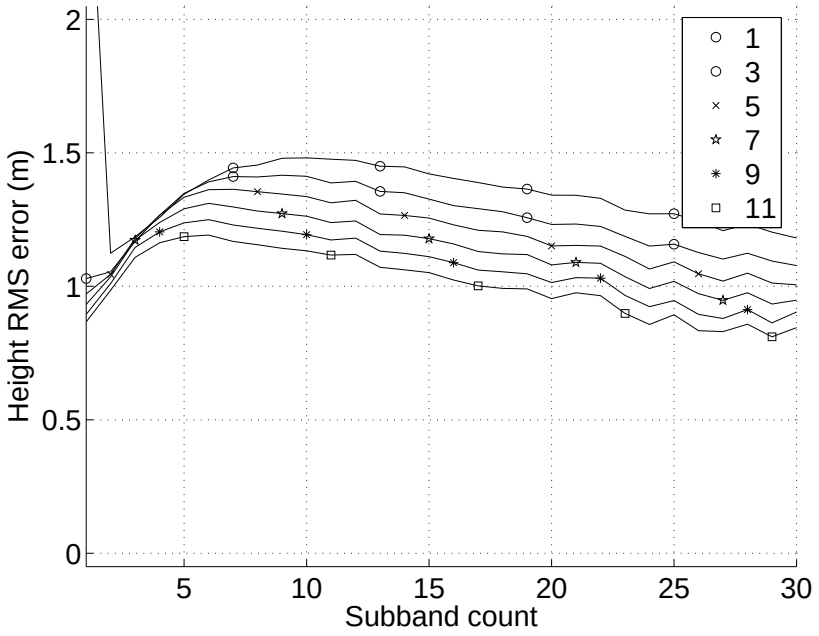


(b) Stuka portion.

Figure 5.16 RMS error of bathymetric height estimates for two portions of the test scene for differing number of looks L; Figure 5.16(a) for the 'background' portion, Figure 5.16(b) for the 'stuka' portion. No additive noise. As the number of sub-bands are increased the RMS error decreases. Increasing the number of along-track looks decreases the RMS error, however the gain diminishes as the number of looks is increased.

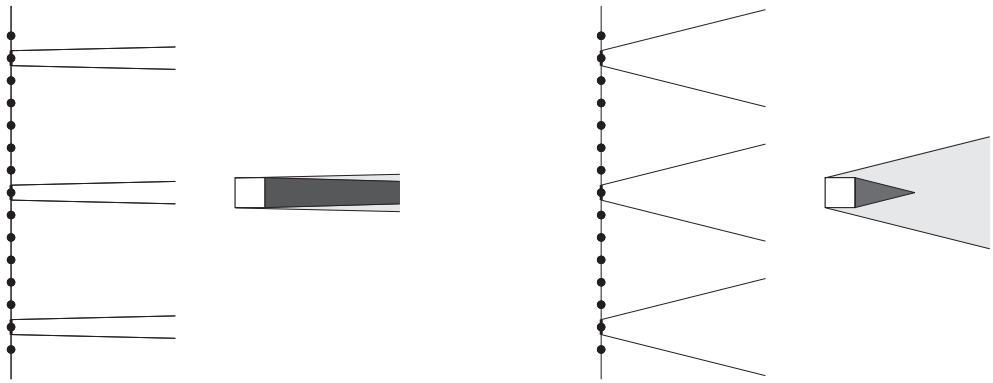


(a) Background portion.



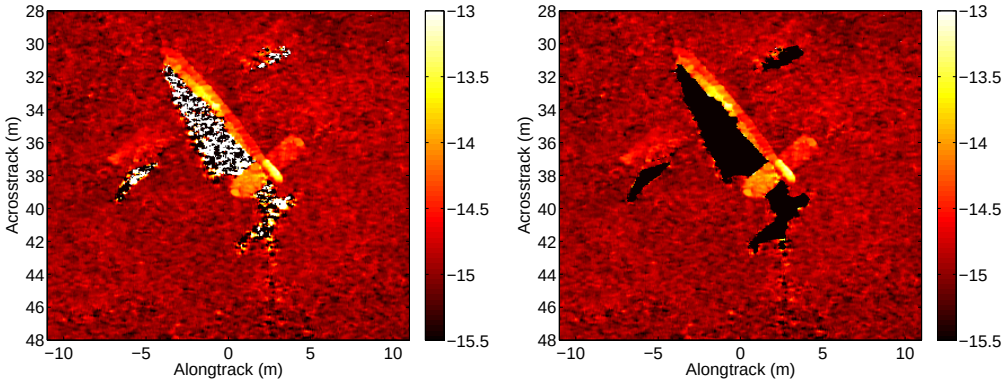
(b) Stuka portion.

Figure 5.17 RMS error of bathymetric height estimates for two portions of the test scene; Figure 5.17(a) for the 'background' portion, Figure 5.17(b) for the 'stuka' portion. Noise added with RMS power 20 dB below the peak of the raw echo response. As the number of sub-bands are increased the RMS error decreases. Increasing the number of along-track looks decreases the RMS error, however the gain diminishes as the number of looks is increased.



(a) Real aperture sonar shadow. (b) Synthetic aperture sonar shadow.

Figure 5.18 Shadow shape with a square object protruding from a flat seafloor for, 5.18(a) a real aperture sonar, and 5.18(b) a synthetic aperture sonar. The shadow consists of two parts, the partially shaded penumbra, and the fully shaded umbra. In the real aperture case the umbra extends much further beyond the object than the synthetic aperture case, while conversely the synthetic aperture case has a much greater penumbra region.



(a) Height before shadow mask applied. (b) Height after shadow mask applied.

Figure 5.19 Height estimates before and after application of a shadow region mask. Shadow mask calculated from the minimum cost for each pixel, thresholded at a coherence of 0.75.

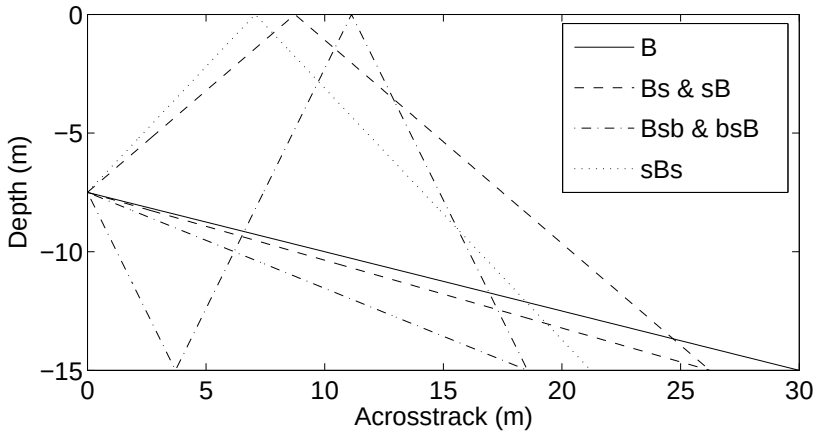


Figure 5.20 Multipath modes for a typical side scanning SAS system. The water column is 15 m in depth with the sonar platform mid-water with a maximum groundplane imaging range of 30 m. The direct path (B) has a slant plane range of 30.9 m, and 5 multipath modes of equal length. Of these paths, two are first order multipath modes (Bs and sB), and three are second order multipath modes (Bsb, bsB, and sBs).

5.7 MULTIPATH

Conventional SAS imaging often assumes the sound energy from the sonar travels in a straight line to the seafloor, and echo return along the same path in reverse. However, in a shallow water environment, the water surface can act as a mirror, reflecting upward travelling sound waves back into the water column. These reflected waves cause targets of closer across-track range appear at the same image range as the desired direct path; a phenomenon known as multipath [Dix and Palmer, 1984]. Higher orders of multipath can also occur from reflections from both the sea-surface and the seafloor, the maximum order restricted by the path length of the direct image path. A multipath naming scheme, as used in [Blanc-Benon and Jauffret, 1997], provides a concise description of the path traveled for both the direct path, and multipath modes. This naming scheme is based on the order and location of each of the 'bounces' of the sound energy. The sea-surface is denoted 'S' and the seafloor denoted 'B'. Later the naming scheme was extended to distinguish between specular and diffuse reflections [Pinto et al., 2004]. For specular reflections (angle of incidence equal to angle of reflection), lowercase characters are used, for diffuse reflections capital characters are used. Thus the direct object path is named 'B', and the two first order multipath modes named 'sB' and 'Bs'. There are two first order modes due to two different directions the sound energy can travel, reflecting from the sea-surface on either the incident or return path. The modes of travel for a typical side scanning SAS system are shown in Figure 5.20.

5.7.1 Resolving multipath modes with a multiple vertical receiver array

Using a traditional interferometer with N_h vertical hydrophones $N_h - 1$ scatters within each resolution cell can be resolved. Thus for a minimalist array of two receivers only one scatter height can be estimated. For resolution cells with multipath modes present the resulting height estimate will depend on the relative phase and strength of the interfering (multipath) signals [Kliger and Olenberger, 1975, Dybdal, 1986], and can appear to have an angle of arrival outside the expected incoming angles [Kraeutner and Bird, 1999]. One solution is to use an interferometric receiver array with a large number of elements, able to resolve the multipath modes from the direct path. However, this is approach is impractical given the physical size of the array required and the constraints of the sonar system [Hayes, 2004].

Alternatively, super-resolution techniques can be used to estimate multiple angles of arrival from within each resolution cell from a small array of hydrophones. Proposed algorithms include Computed Angle-of-Arrival Transient Imaging (CAATI) [Kraeutner and Bird, 1999], Coherent Source Direction Estimation (CSDE) [Xu and Stewart, 1999], and Bayesian techniques based on the RELAX algorithm [Li and Stoica, 1999, Hayes, 2004].

The RELAX algorithm attempts to estimate multiple angles of arrival in a round-robin fashion. After estimating the dominant angle of arrival, this component is subtracted from the original signal. A second estimate is then made from the residual signal, again estimating the dominant angle of arrival. After repeating this process for the number of multipath modes expected the process is iterated, with each estimation made with the estimates of the remaining angles of arrival subtracted. The technique has shown good results on simulated data with a single sea-surface multipath mode present [Hayes, 2004].

5.7.2 Beam-steering to reduce multipath modes

Many of the multipath paths are transmitted or received by the sonar at angles above horizontal, while the desired direct echo is almost always below horizontal. Thus if the transmitter and receiver element beam-patterns are adjusted to have high attenuation at these above horizontal angles, many of the multipath modes can be removed. The adjustment of the beam-patterns can be achieved in several ways; altering the shape of the elements, acoustic shading with a physical waveguide, or synthetically through the use of a vertical array of small elements. Of these methods, the vertical array is the most flexible, allowing different beam-patterns to be used by post-processing the data. However, this method requires a substantial increase in towfish hardware costs, requiring large transmitter and/or receiver vertical arrays.

Transmitter vertical beam-pattern

Often in a InSAS system, the transmitter is constructed of a grid of smaller transmitter elements, allowing for multiple power-amplifiers to increase acoustic power.[9] Traditionally these elements are driven simultaneously with identical transmit waveforms, giving a beam-pattern shape dictated by the overall dimensions of the array. More recently, advances in digital electronics allow the capability to drive each element of the array with independent signals, allowing the overall beam-pattern to be altered. The extra towfish hardware and power required to drive the transmitter elements individually is often negligible.

Receiver hydrophone beam-pattern

In order to increase SAS mapping rates, many recent SAS towfish designs have employed large along-track receiver arrays, normally with just one element in the vertical direction. To provide some synthetic vertical beam-pattern control, a vertical array of these large along-track arrays would be needed. This is is an expensive prospect, both for the actual hydrophones and associated electronics, and the increase in data storage requirements.

5.7.3 Multipath shadow degradation

The shadows considered in Section 5.6.7 are those from a direct imaging path, both in the transmitting and receiving directions. In addition to this direct imaging mode, many multipath imaging paths are present. These multipath modes occur due to the sound energy being reflected from both the seafloor and the sea-surface. An example of multipath is shown in Figure 5.20 for a single ping. Higher order multipath modes can also exist when more than one ping is considered. In this case the energy from the previous pings also have equal path lengths.

The resulting image is therefore the sum of all direct and multipath path lengths of equal length. In a shallow water environment there are several orders of multipath of equal length to the direct path.

With an InSAS system with N_h vertically separated hydrophones, $N_h - 1$ incoming angles of arrival can be resolved. However, the angular resolution of a small array is poor, so the incoming wavefronts to be resolved must be separated by a large angle. It is therefore impossible to remove any multipath modes with an angle of arrival similar to that of the direct echo.

[9] For example, the KiwiSAS transmitter consists of an array of 12 horizontal by 3 vertical elements.

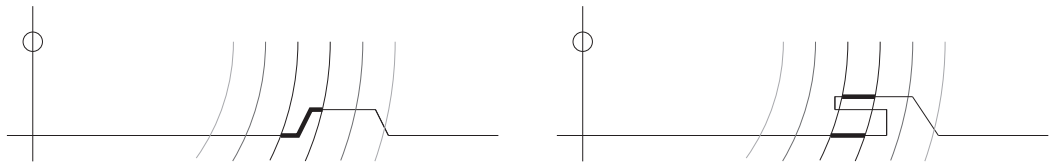


Figure 5.21 Layover occurs for objects with near or beyond vertical front faces. Figure 5.21(a) shows a near vertical face with multiple reflector heights within the same resolution cell. In the more extreme case of an overhanging object, multiple distinct surfaces can be present within a resolution cell, as shown in Figure 5.21(b).

5.8 LAYOVER

Layover occurs when there is a significant height change of the seafloor within a single across-track resolution cell. In general, layover only occurs for the front face (as viewed by the sonar platform) of an object, since the rear face will often be in a shadow region. This steep front face can occur as part of the seafloor surface, for example a rocky cliff, or because of a proud object sitting on the seafloor. An example of this form of layover is shown in Figure 5.21(a). In the extreme case for an overhanging sea-surface or object, there can be multiple distinct surfaces within a resolution cell, as shown in Figure 5.21(b). Overhanging structures are unlikely to be present in the seafloor itself (although could occur because of an overhanging cliff), more likely present because of a man-made object on the seafloor.

When imaged by a side-scanning SAS system, the resulting echo response for a target with significant height variation will be the sum of each of the surface segments within the resolution cell. Because of this summing process, the conventional assumption for InSAS bathymetry of a single reflector within each resolution cell is violated. By estimating the scene height using the (incorrect) assumption of one reflector per resolution cell, the result is an average height of all the reflectors in the resolution cell. The average reconstructed height is not necessarily the same as the average of the physical scene heights; it is an average of the physical scene heights weighted by their reflected signal strength, resulting in an overall height biased towards the stronger reflector [Kliger and Olenberger, 1975, Dybdal, 1986, Hayes et al., 2005]. For example a sharp cornered, man-made object proud of a flat seafloor will have an reconstructed height biased towards the strongly reflecting man-made object.

Another distortion of proud objects is that of foreshortening. When the scene is viewed from an angle, the range to the top of a proud object is less than its true groundplane range. Consequently, the top of the object appears closer to the imaging platform in the bathymetric heightmap. This phenomenon is more significant in InSAR applications because of the steeper look angle of most InSAR platforms, and the severe height changes of some Earth structures (for example a mountain range). For a

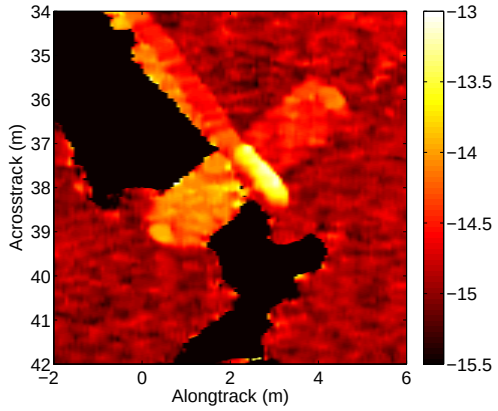


Figure 5.22 Layover occurring in resolution cells with multiple reflectors. Scene is the tail section of an aircraft with both tail-planes located approximately 1.2 m above the nominally 15 m deep seafloor. The right tailplane has a incorrect reconstructed height, midway between the correct tailplane height and the sea-surface. Layover is the cause of this height error, occurring because each resolution cell has two reflecting surfaces, the tailplane and the seafloor. The left tailplane does not exhibit this layover error since the plane fuselage is shadowing the seafloor, so each resolution cell only contains one reflecting surface of the tailplane. The vertical tail section exhibits foreshortening; the top of the vertical section appearing closer to the sonar platform than the vertically aligned but lower fuselage.

typical SAS platform, the viewing angle of the scene is often only a few degrees below horizontal, minimising the foreshortening distortion. An example of foreshortening is shown in Figure 5.22 where two vertically aligned, but separated, points in the scene are misaligned in the bathymetric heightmap.

5.9 KIWISAS SEA-TRIAL RESULTS

Sea-trials were conducted using the KiwiSAS-IV sonar system in late January and mid February 2006 with the sonar configured with a vertical array of three hydrophones. For more detail on the KiwiSAS-IV system, see Section B.5 and [Hayes et al., 2002, Hayes, 2003]. The sea-trials were conducted in Lyttelton harbour in the vicinity of Parson's Rock, a volcanic protuberance from an otherwise bland flat mud seafloor (see Figure 5.23). The harbour in this area is 10-12 m deep, with the top of Parson's Rock extending approximately 6-7 m above the harbour floor. The sonar was operated approximately mid-water using a depressor chain, nose-towed with a forward velocity of approximately 1 m/s. The echoes from each hydrophone were recorded in raw form to an on-board hard-disk, later pulse-compressed in range, and in azimuth using the wavenumber algorithm (see Section 1.1.4). Backscatter intensity imagery for a small portion of this sea-trial is shown in Figure 5.24. Distinct differences between the low (30 kHz) and high (100 kHz) bands are evident in the backscatter intensities of Figure 5.24. These differences can be partially explained by the composite transmitter/hydrophones beam-pattern, the high band having a narrower beam-pattern. This explains the amplitude differences at ranges less than 20 m, however, significant unex-

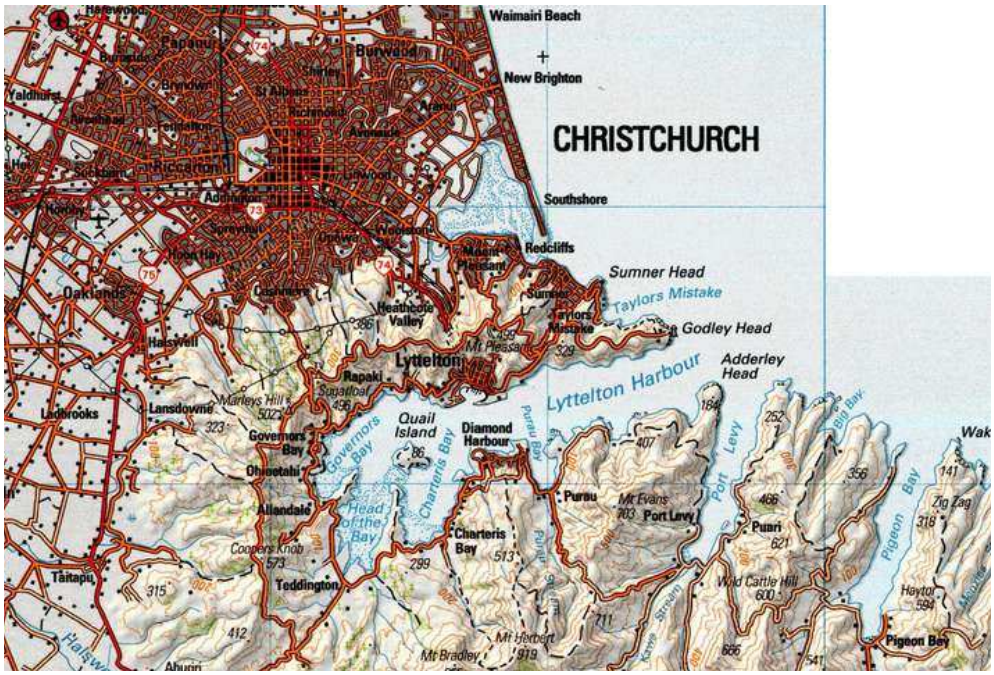
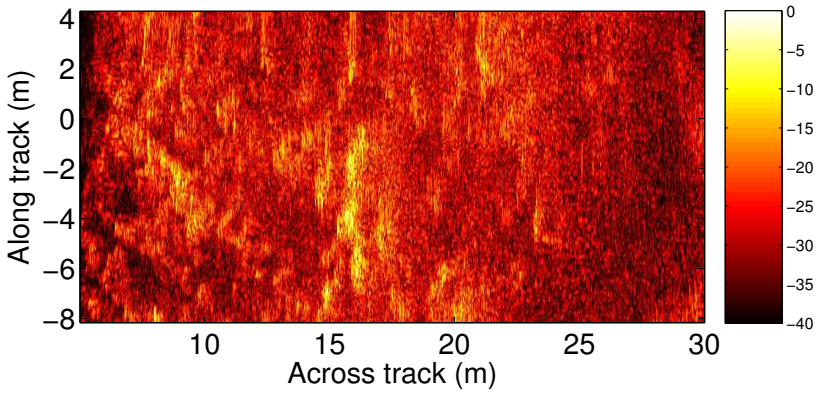


Figure 5.23 Lyttelton Harbour is the primary survey site for most of the KiwiSAS sea-trials. Lyttelton Harbour is located South-East of Christchurch, New Zealand. Water column is approximately 10–12 m in depth, with a nominally smooth, muddy sea-floor. Sea-trials are often performed near 'Parson's Rock', a volcanic protuberance near the inlet to Purau Bay.

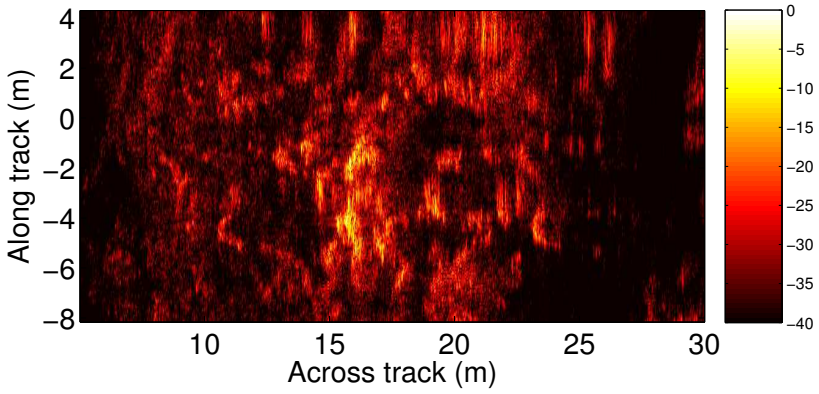
plainable differences also exist. These differences suggest a frequency dependence on the backscatter intensity, possibly indicating differences in acoustic penetration into the seafloor [Gough et al., 2006].

For coherence estimation, the data is first sub-banded into several narrow-bands. This decreases the effect of the footprint shift between the receivers, increasing the observed coherence between the signals. Using just the central narrowband portion, the coherence of Figure 5.25(a) and 5.25(b) is obtained. For the coherence calculations, just the top and middle receiver channels are used. The coherence can be seen to be higher for the lower frequency band compared to the higher frequency band. This frequency dependence is expected; any small geometric error will have more effect on the phase of the higher band.

Figure 5.25 also shows a direct interferogram between the upper and middle receiver channels, calculated as an average over a small region (9 in range, 5 in azimuth) of image pixels. The phase differences in both the low and high bands are approximately zero for an across-track range of about 28 m indicating the seafloor is at a depth similar to that of the towfish. This is consistent with previous depth-soundings of the area performed earlier, indicating the presence of a large "sea mound" in this area. Most likely this object is a large rock protrusion covered with a layer of mud and silt, and possibly areas of seaweed. From the point of zero phase at 28 m, the low band exhibits

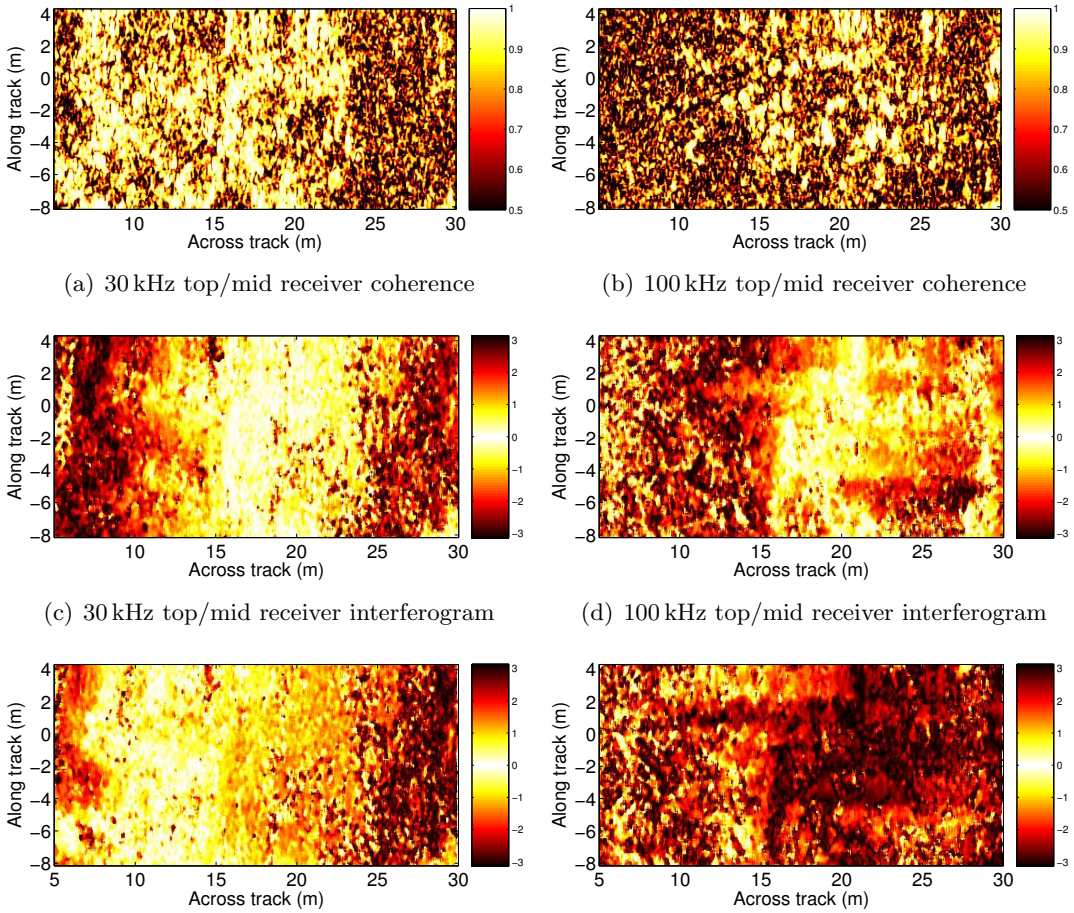


(a) 30 kHz top receiver intensity (dB)



(b) 100 kHz top receiver intensity (dB)

Figure 5.24 Sea-trial backscatter imagery after synthetic aperture processing. (a) and (b) show the top receiver echo intensity from each band.



(a) 30 kHz top/mid receiver coherence

(b) 100 kHz top/mid receiver coherence

(c) 30 kHz top/mid receiver interferogram

(d) 100 kHz top/mid receiver interferogram

(e) 30 kHz top/mid receiver remapped interferogram

(f) 100 kHz top/mid receiver remapped interferogram

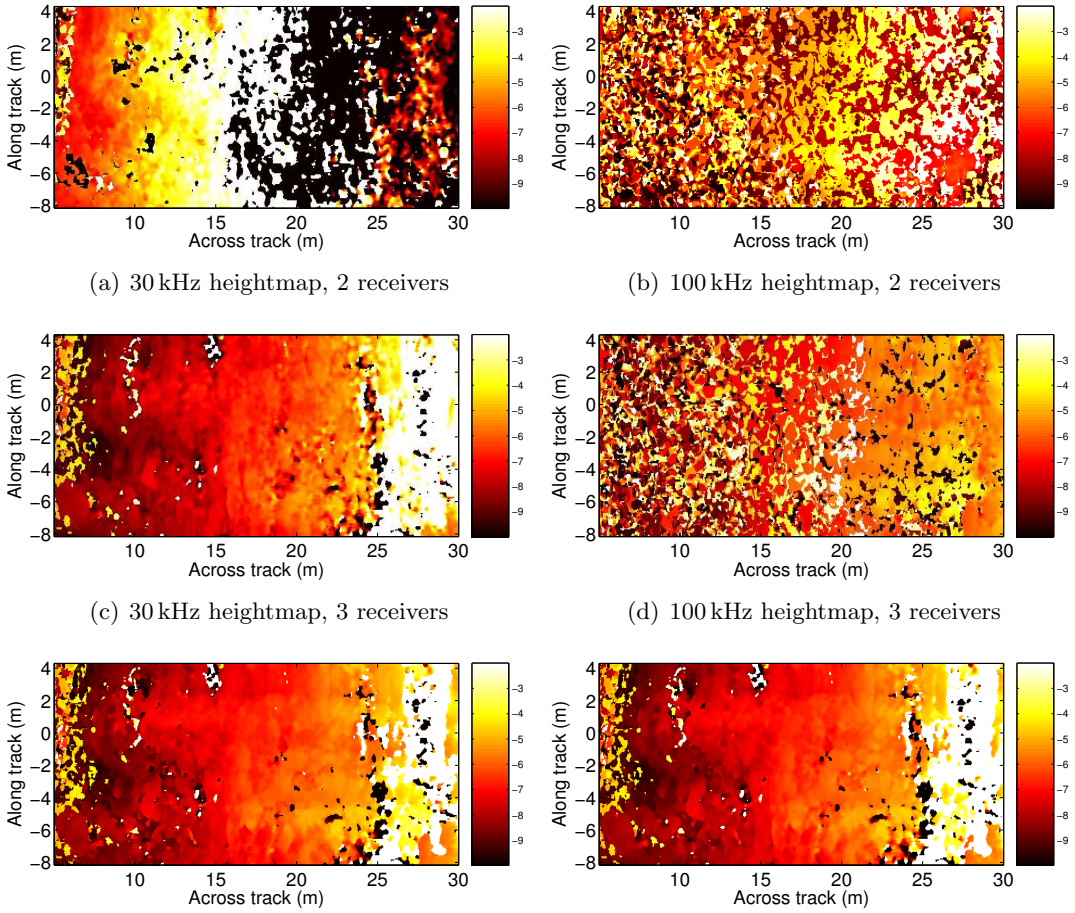
Figure 5.25 Sea-trial results. (a) and (b) top receiver echo intensity from each band, processed using the wavenumber algorithm. (c) and (d) coherence between top and middle receivers. (e) and (f) interferogram between top and middle receivers. (g) and (h) interferogram between top and middle receivers after remapping to a groundplane of -9 m below the sea-surface.

a smoothly changing phase difference into about 13 m; closer than this showing many areas of phase wrapping. Due to the higher frequency of the high band, the region of smooth phase gradients is much smaller, only extending into a range of approximately 25 m from the full range of 30 m. The extent of this smooth phase difference gradient can be extended by remapping the data onto a common groundplane closer to the actual seafloor height. The remapping of the echoes to a common groundplane closer to the actual seafloor height stretches and aligns the echoes such that a common point in the original scene will occur in the same range position in each of the echoes, improving the signal registration and hence higher coherence between the echoes. The effect of this remapping can be seen in the last two images of Figure 5.25 where the area of smooth phase difference gradient has been extending to closer across-track ranges.

The seafloor boundary between the water column and the mud/silt is most likely not a distinct transition, more likely is a gradual transition with increasing concentrations of solid material. This transition will cause volumetric scattering, more predominant in the higher (100 kHz) band.

Height reconstructions have been produced from the dataset using various combinations of the two lower and upper datasets, shown in Figure 5.26. Processing the lower band alone provides a smooth height estimate over a major portion of the scene with regions of inconsistency at near ranges. As expected when the upper band is processed alone, significant errors occur at close ranges, consistent with the coherence results earlier. Comparing the results from the two individually processed bands shows general agreement at far ranges, suggesting no phase wrapping has occurred in either case.

Combining the two frequency bands into one overall estimate should provide a more accurate height estimate due to the increase in scene looks. The height map using both frequency bands is shown in Figure 5.26(e). However, the combination of the two bands is complicated by amplitude differences between the two bands. If one frequency band has a significantly stronger signal strength the composite cost function formed within the ML estimation process will be dominated by this higher amplitude signal. This will cause the overall ML estimate to be biased towards this frequency band. The amplitude differences between the frequency bands are caused by many factors such as transmitter and receiver sensitivities, scene reflectivity differences, and absorption within the medium. Since these cannot be fully characterised the two frequency bands must be scaled before combination in the ML process. As a first approach the bands are normalised to have the same peak intensity within the scene, the resulting height shown in Figure 5.26(f). Ideally the bands should be weighted by the expected variance of the signal.



(a) 30 kHz heightmap, 2 receivers

(b) 100 kHz heightmap, 2 receivers

(c) 30 kHz heightmap, 3 receivers

(d) 100 kHz heightmap, 3 receivers

(e) 30 kHz and 100 kHz heightmap, 3 receivers

(f) 30 kHz and 100 kHz heightmap, 3 receivers, normalised

Figure 5.26 Heightmap reconstructions from sea-trial results. (a) and (b) 2 receivers combined for each band respectively, (c) and (d) 3 receivers combined for each band respectively, (e) 3 receivers, 2 frequency bands, and (f) 3 receivers, 2 frequency bands after normalisation between the bands.

Chapter 6

CONCLUSIONS AND SUGGESTED FURTHER WORK

Broadband InSAS produces poor height estimates if interferometric techniques are applied on a pixel-by-pixel basis. For interferometric baseline lengths of greater than the across-track resolution cell, footprint shift between the receiver channels causes per-pixel phase comparisons to be meaningless.

Ensuring high coherence between the receiver channels is essential for high accuracy height estimates. Even small losses in coherence from the ideal unity cause a large increase in the error of the height estimate. When the overall coherence is less than unity, significant gains in the height accuracy can be made by producing multiple 'looks' of the scene. Combination of these looks into one height estimate will reduce the error in the height estimate, although at the expense of a lower ground-plane resolution in some cases. For typical coherence levels of a practical InSAS systems, a large number of looks is required to provide low variance height estimates.

The coherence between receiver channels is a combined metric composed of five different coherence components, each caused by a different mechanism. Of the five coherence components, the most dominant for InSAS applications is that of footprint shift. However, footprint shift can be corrected by resampling the data to the correct height. The difficulty here is the selection of a ground-plane to re-sample onto. For ideal correction, the ground-plane height must be known with a resolution equal to the backscatter resolution. If such a height-map is known, there is no need to perform interferometry since the answer is already available. However, partial correction to a lower resolution ground-plane still gives substantial coherence gains, improving the height estimation accuracy. Some coherence components cannot be corrected, limiting maximum achievable coherence, even after correction for footprint misalignment.

Sonar design can alter coherence across the interferometric array significantly. One parameter shown to be of high importance is that of the along-track ambiguity to signal ratio (AASR). For every point in the scene, synthetic aperture processing causes the generation of 'grating lobes'; large areas of low level interference at significant along-track offset. The grating lobes combined from every point within the scene can be considered a source of noise, the level of which is defined by the AASR. For a marginally

sampled system with $D/2$ along-track sample spacing, the level of background noise given by the AASR is in the order of -13 dB, of low consequence for intensity based backscatter imagery. However, when the phase of the signal is considered for interferometric applications, corruption from such values of AASR cause significant losses in coherence by a factor of 0.9523. By increasing the along-track sampling rate, significant increases in AASR can be made; $D/3$ sampling yielding an AASR of -24 dB. Significant gains can also be made through careful selection of transmitter and receiver element sizes; the combined beam-pattern controlling the energy spread into the grating lobes, and ultimately the coherence across the interferometric receiver array. For example, a system with a transmitter element of $1.6\times$ greater in size than receiver elements, the AASR is improved to -23 dB for the $D/2$ sampling case. A combination of increased along-track sampling and unequal transmitter and receiver sizes can yield AASR values of less than -30 dB, equivalent to a coherence level of 0.9990. During the design of an InSAS system, it is recommended to carefully consider the along-track size ratio of the transmitter and receiver elements in order to minimise the detrimental effect of grating lobes. Practical considerations such as minimum element sizes may impose a limit on the sonar design choices, but significant gains in AASR can still be made with subtle changes in element ratio. Since grating lobes are a by-product of synthetic aperture processing, the associated coherence loss from AASR is not seen for real-aperture interferometric sonar systems.

To reduce the coherence loss from footprint shift across an interferometric receiver pair, the data can be resampled to a common groundplane, ideally to a surface as close to the actual height as possible. To form this initial lower-resolution surface, the imagery from each receiver of the interferometric array can be back-projected to a range of ground-planes, through which a surface can be fitted using a technique called belief propagation. This method incorporates prior knowledge of the expected seafloor terrain rates of change. After finding a lower resolution surface from belief propagation, the data can be corrected to this surface and then be processed using ML techniques to provide the final high resolution height-map. Although belief propagation could be used to find a final height estimate of the scene, the computational load is high. Also, the use of a ML estimator yields a height estimate over a continuous height variable, rather than the discrete levels produced from belief propagation.

Maximum likelihood techniques can be used to detect areas of shadow. Within the areas of shadow, the received waveform should be non-existent. As a result, each look generated is a sample of noise containing no scene reflection. When these noise samples are combined within the maximum likelihood estimator there is no coherent gain at a particular solution. As a result, the cost function minimum will be higher in level than for a non-shadowed scene position. By thresholding the searched minimum cost function at each pixel location across the scene, areas likely to be in shadow can be detected. The detection of shadow regions could provide a useful source of information

for subsequent CAD/CAC processing.

Multipath can cause errors in the height estimates. Typical shallow-water InSAS geometry produces many modes of multipath, violating the central assumption of only one target per resolution cell. Multipath modes could be rejected by use of a larger vertical array of hydrophones, allowing synthetic beam-forming in the vertical direction. However, the vertical resolution required to isolate some modes of arrival would require an impractically large vertical array.

6.1 FURTHER WORK IDEAS

Sea-surface multipath rejection The presence of multipath echo returns from the sea-surface distinguishes InSAS from InSAR; no analogue is present in InSAR systems. The effect of multi-path is a severe reduction in signal coherence across the interferometric array, causing substantial height errors in the bathymetric height estimation. Due to poor vertical resolution, multi-path modes with angles-of-arrival close to that of the direct signal return cannot be rejected with traditional receiver beam-forming approaches. Instead, it is important to reduce the multi-path signal amplitude using physical design parameters of the sonar system. This reduction of multi-path could take the form of vertically tapered transmit signals, acoustic shields/baffles, coded waveforms, etc.

Incorporate footprint shift correction into the ML estimate search During the maximum likelihood estimation search the seafloor height is indirectly used as the search parameter (search performed over time-difference between the receivers). At each iteration, the echo data could be re-sampled onto the seafloor surface of the previous iteration. This should minimise the effect of coherence loss from footprint shift, corrected to the most recent height estimate. However, the computation load of performing the remapping at every iteration may prove too expensive for the gain in coherence.

Global searching of cost functions Normally the seafloor height from a scene is estimated on a point-by-point basis. Each point of the scene has a height calculated based on the echo returns. As an alternative, the seafloor height could be estimated over a larger patch of seafloor. The maximum likelihood framework presented in Chapter 4 could be extended to directly estimate the entire surface of the scene in one step, incorporating the data from all points within the scene. Prior information of the scene height statistics would also be required. However, the computational complexity of a globalised approach would be prohibitive. Instead, approaches such as belief propagation (see Appendix A) could be extended to find a global height surface.

Volumetric scattering from seafloor penetration. In many circumstances, the transition at the seafloor from water to substrate will not be a defined boundary. In situations with a more gradual transition, volumetric scattering from within the transition zone will occur. Volumetric scattering violates the assumption of a single scatterer position within each resolution cell.

Repeat-pass interferometry. If the position of the sonar system could be estimated to sub-wavelength accuracy, repeat-pass interferometry may be possible. With sufficient positioning accuracy, the imagery from the multiple passes could be registered and formed as an interferometric set. As an intermediate goal, differential interferometry could be attempted; differences in scene topography between interferometric imaging could be used as a measure of seafloor changes over a longer period of time.

Parametric modeling of seafloor Rather than define the height of the seafloor as a matrix of height values, a parametric model of the surface could be formed, for example, using radial basis functions. With careful selection of a parametric model, erroneous rapidly changing portions of the seafloor would be smoothed. For areas of more complicated seafloor structure, such as targets proud of the seafloor, separate parametric models could be used to describe this local area. Fitting of parametric models to localised objects of interest could be incorporated into CAD/CAC type algorithms.

Appendix A

RECONSTRUCTING SEAFLOOR BATHYMETRY WITH A MULTICHANNEL BROADBAND INSAS USING BELIEF PROPAGATION

This appendix is a reproduction of the publication "Reconstructing seafloor bathymetry with a multichannel broadband InSAS using Belief Propagation", by Philip J. Barclay, Chris J. Forne, Michael P. Hayes, and Peter T. Gough, appearing in the OCEANS 2003 Proceedings, Volume 4, Pages:2149 - 2154 Vol.4, 22-26 September 2003.

Reconstructing seafloor bathymetry with a multichannel broadband InSAS using Belief Propagation

P. J. Barclay, C. J. Forne, M. P. Hayes, and P. T. Gough

Acoustics Research Group,
Dept. of Electrical and Computer Engineering,
University of Canterbury,
Private Bag 4800, Christchurch, New Zealand

{*p.barclay,c.forne,m.hayes,p.gough*}*@elec.canterbury.ac.nz*

***Abstract*— Synthetic Aperture Sonar (SAS) provides high resolution imagery with range independent resolution. However, imagery has been limited to two-dimensions, showing acoustic reflectivity for a given scene. To further enhance the usefulness of a SAS system this needs to be combined with a topographical height map of the same scene. This paper presents a novel technique to obtain the third-dimension by first remapping on to a set of common ground-planes, with the most likely surface through this volume found using Belief Propagation. This technique allows any a priori information to be incorporated and thus provides a maximum a posteriori (MAP) estimate of the seafloor height. Results are presented from a rough seafloor simulator with height reconstructions using this technique.**

I. Introduction

Synthetic aperture sonar (SAS) is similar to synthetic aperture radar (SAR) in that it can generate high-resolution imagery independent of range. This is achieved through coherent summation of successive echo signals to synthesise an aperture many times longer than the towfish [1]. While the synthetic aperture technique generates high-resolution imagery, it does not provide bathymetric information regarding the topology of the seafloor. A solution to this is to use two or more vertically separated receivers. By measuring the delay differences between the displaced receivers, the angle of arrival of the incident echo wavefront can be estimated. Then combining the angle estimates with the range measurements, an estimate of the height of the scatterers can be derived.

While interferometric aperture synthesis has been successfully achieved in radar (InSAR) [2]–[4], the sonar equivalent is complicated by the slow speed of sound necessitating wider beamwidths and broader bandwidths, by unwanted reflections from the sea-surface, and by greater unknown motion errors.

The technique presented in this paper employs a statistical filtering approach to search a volume of height likelihoods. This volume is generated by remapping the slant-range imagery onto a set of height differing ground-planes. The approach eliminates the need for 2-D phase unwrapping required by direct interferometry techniques.

The imagery presented in Sec. (IV) is obtained from simulations of the KiwiSAS-IV system developed at the University of Canterbury. This sonar system operates simultaneously at two frequency bands centered on 30 kHz and 100 kHz, each with a bandwidth of 20 kHz. The system has one projector, and three vertically separated hydrophones simultaneously sampled. For more details of this system see [5], [6].

II. Direct Interferometry

An interferogram is formed from two complex images (q_1), and (q_2) by taking the Hermitian product of the two complex images

$$\chi = q_1 q_2^* \tag{1}$$

where $(^*)$ denotes the complex conjugate. The interferogram is then simply the phase of χ

$$\hat{\psi} = \angle\chi \tag{2}$$

The interferometric phase can then be used to estimate the height of the imaged slant range point. This can be calculated for every point in the images, resulting in an overall three-dimensional topography estimate. This technique is well established in both airborne and spaceborne Synthetic Aperture Radar (SAR) systems [3], and more recently Synthetic Aperture Sonar (SAS) [7]–[10].

Due to the specular nature of SAS imagery, any shift in imaging geometry causes a decorrelation [11]. Since interferometry uses two different imaging looks of the scene, decorrelation effects can be observed. This has the effect of producing speckly interferograms, with largely altered phase differences for a small change in imaging geometry. If these phase differences are used directly to produce a height estimate of the original scatterer, unreasonable height variations can occur.

The interferometric phase difference is also calculated modulo (2π) requiring two-dimensional phase unwrapping [12] in all but a few special interferometric applications [13].

III. Bayesian Height Estimation

Some SAR researches have measured the coherence between pairs of ground-plane remapped data to estimate the correct depth [3]. However for shallow water SAS the coherence only varies slightly with seafloor depth error resulting in low height resolution.

In this paper we estimate the seafloor height be remapping the data over many different heights to create a volume of hydrophone variances. This volume is then searched for a minimum to give a seafloor surface estimate. This approach works with an arbitrary number of hydrophones and frequency bands and is a maximum likelihood (ML) estimator. However due to residual decorrelations in the data, the height estimates are noisy so we then statistically filter these estimates including a prior information about the expected seafloor topography. This then provides a maximum a posteriori (MAP) estimate of the seafloor height.

This process can be divided into four steps:

1) Generate volume of ground-plane remapped images.
2) Calculate variance of each voxel.
3) Convert voxel variance to likelihoods.
4) Find MAP estimate of height using Belief Propagation.

A. Groundplane Remapping

During the Synthetic Aperture reconstruction process the third dimension (height) of the scene is not used. Instead the height is assumed to be zero, and all reconstruction is done on the slant-range plane. This has the effect of producing an increasing acrosstrack shift in the image for increasing height difference between the sonar and the seafloor. For an interferometric setup, the hydrophones are vertically separated but all share a common transmitter array. Thus the acoustic path difference is dominated by the acoustic receive path only.

To compensate for the differing receiver path lengths, groundplane remapping can be performed to resample the data from the various slant range planes to a common groundplane. To do this remapping the topography of the seafloor must be known. Since this height information is not known initially an estimate must be made. From this initial height estimate the true distance to a given acrosstrack point can be found, then used as an index into the original slant range plane data. Subsample interpolation must then be used to estimate the signal at the corrected range. A phase correction term is also applied corresponding to the phase change of the original modulated signal.

For each receiver R_n the geometry is as shown in Fig. 1 where s is the slant-range data, and g the remapped ground-plane data. The envelope remapping is then given by the equations:

$$\begin{aligned} s_n(x) &= \sqrt{g(x)^2 + \left(g(z) - R_n(z)\right)^2} \\ s_n(y) &= g(y) \end{aligned} \tag{3}$$

The phase correction term required to compensate for the demodulation of the original acoustic signal is then

$$\varphi = \exp\left(\frac{-j4\pi f_c s(x)}{c}\right) \tag{4}$$

where f_c is the center frequency, (c) is the speed of sound in water.

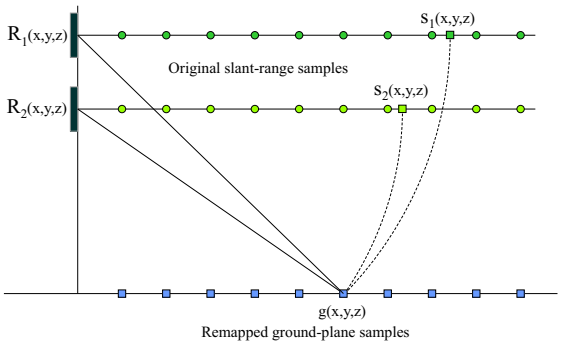


Fig. 1. Remapping slant-range plane to ground-plane.

The overall effect of this process can be seen in Fig.2, for the case of a single point reflector located 45 m directly acrosstrack, 5 m below the sonar. Notice the peaks of the original slant range curves are centered at approximately 45.2 m, and after remapping to the correct groundplane they are centered on the original simulation position of 45 m.

For a flat seafloor scene, ground-plane remapping can be seen by comparing the interferograms of the original slant-range plane data, and after remapping to the 'correct' ground-plane as shown in Fig.3. The original interferogram is very speckly in nature, and shows several interference fringes in range across the image. After remapping to the 'correct' height-plane the speckle has been reduced, and the interference fringes have been removed leaving a mean phase difference of zero. The speckle reduction is as a result of decreasing the decorrelation between the scenes with corresponding pixels footprints now aligned. The remaining speckle is from the slightly different angle of viewing of the scene from the two receivers [14]. Bands of higher speckle can be seen across the interferogram at ranges of 23 m, 11 m, 6 m, and several < 5 m corresponding to the nulls in the vertical beam pattern of the transducers.

B. Calculating Variance

The variance between the remapped images is calculated using

$$v = \sum \frac{\left|\hat{q}_n - \mu\right|^2}{n-1} \tag{5}$$

where μ is the mean of $\hat{q}_n$ and n is the number of receivers.

C. Conversion to Likelihoods

Each remapped pixel on each ground-plane is converted to a likelihood. This likelihood gives a measure of how likely the pixel has been correctly remapped. To do this the samples are assumed to be from a normal distribution, so the likelihood L

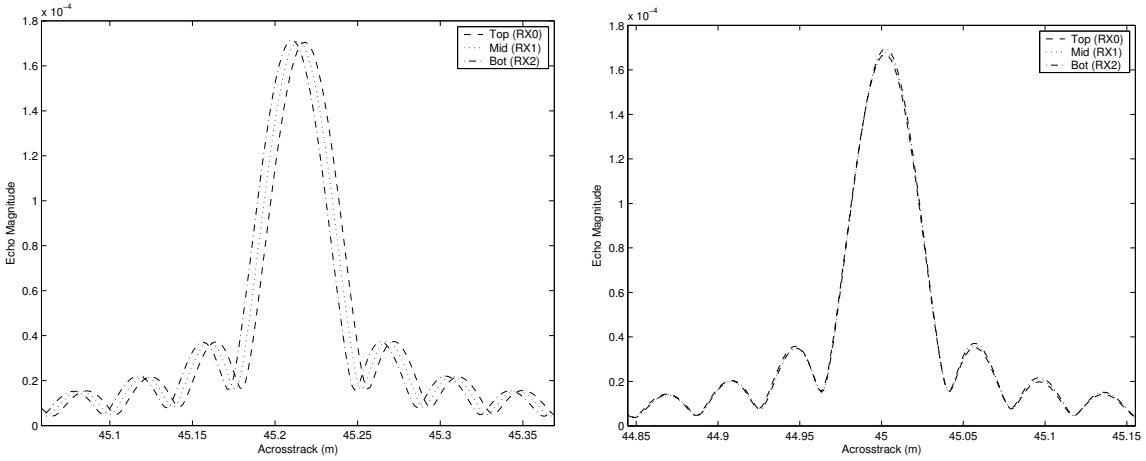


Fig. 2. Signal envelope of a single point target at 40 m range, 5 m below sonar. (a) reconstructed on slant range plane, and (b) remapped to correct groundplane.

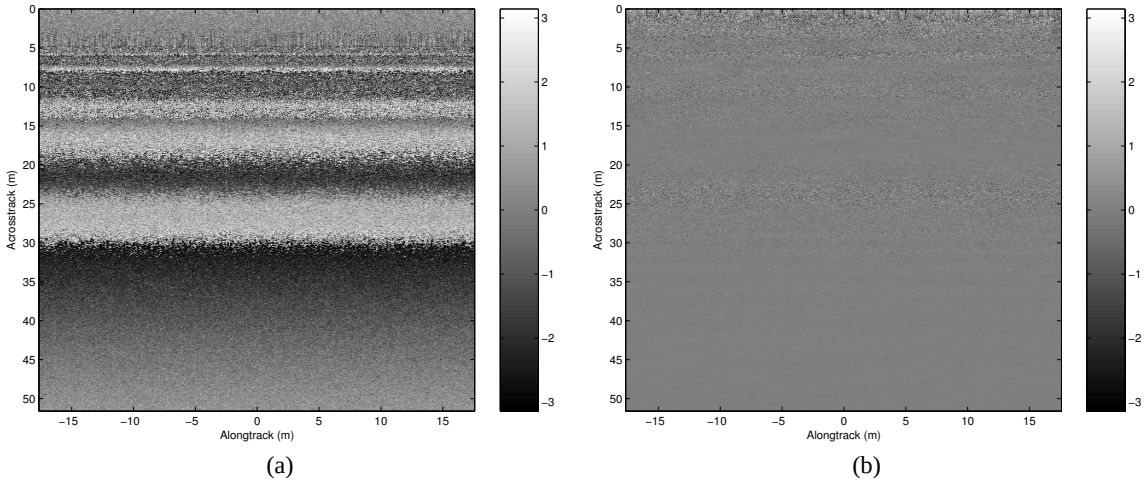


Fig. 3. Interferogram for (a) slant-range plane data, and (b) ground-plane data remapped to the 'correct' height. Receiver separation 0.15 m.

at pixel i for the h^{th} height can be derived from the variance v of the complex samples using

$$L = \frac{1}{\sqrt{2\pi a}} \exp\left(-\frac{v^2}{2a^2}\right) \tag{6}$$

where a is the variance of the distribution. This relationship relies on the remapped vectors having low variance when remapped onto the correct ground-plane. An initial height map can be obtained from this likelihood volume by taking the maximum likelihood for each pixel.

D. Belief Propagation

Belief Propagation is an efficient technique based on local message passing can be used to find the most likely state of a system given any available evidence. By appropriately modeling the system, this approach can be applied to bathymetric data to produce an improved estimate the sea floor. Although originally proposed by Pearl [15] for performing probabilistic reasoning on Bayesian networks, the technique and has since been successfully applied to the stereo imaging field [16]–[18]. In most of this work, the system is modeled as a hidden Markov Random Field (MRF). Belief Propagation is then applied to this model, acting as a statistical filter to find the most likely surface, given a volume of measured likelihood estimates and the expected variation between neighbouring points.

To apply this technique to bathymetric data, the scene is represented as a connected 2D array of nodes, whose value corresponds to surface height. Associated with each node are a set of beliefs which give a measure how likely the surface is to be at any particular height. The objective is to calculate these beliefs as accurately as possible, by propagating the original data throughout the network.

Using the "max-product" algorithm [19], the belief B at a node i for the h^{th} height at the n^{th} iteration is computed using

$$B_i^n(h) = L_i(h) \prod_{j \in N(i)} M_{ji}^n(h), \tag{7}$$

where $N(i)$ is the set of neighbouring nodes to node i, $L_i(h)$ is the initial likelihood computed from the measurement data, and where $M_{ji}^n(h)$ is the message from node j to node i at iteration n, computed by

$$M_{ji}^n(h) = \max_{h'} \left(\Psi(h, h') \frac{B_j^n(h')}{M_{ij}^{n-1}(h')} \right). \tag{8}$$

Here $\Psi(h, h')$ is a compatibility function that describes the joint probability distribution of neighbouring node heights, and $M_{ij}^{n-1}(h)$ is the message from node i to node j from the previous iteration.

In practice the messages are normalised, and a time average operation is performed to stabilise the algorithm and prevent local oscillations

$$M_{ij}^n(h) = \frac{M_{ij}^{n-1}(h) + M_{ij}^n(h)}{\sum_{h=1}^{H} \left(M_{ij}^{n-1}(h) + M_{ij}^n(h) \right)} \tag{9}$$

where H is the number of height measurements and the initial messages are given an identical weighting,

$$M_{ji}^0(h) = H^{-1}. \tag{10}$$

The algorithm runs iteratively until it converges to a steady solution after approximately 40 iterations. The most likely surface is given by the height at each node with the highest belief value. In other words,

$$\hat{h}_i = \max_h B_i(h). \tag{11}$$

IV. Imagery

The simulations are of a rough seafloor using a Monte Carlo integration of many simple point scatterer responses. This model neglects occlusion and multiple scattering but gives a reasonable approximation of the coherent speckle effects that occur with a real sonar. The rough seafloor was modeled as a fractal surface with a Goff-Jordan power spectrum [20] using an approximate Fourier method with oversampling. The 'man-made' objects are then introduced by modifying the reflector heights using an analytical terrain description.

Fig. 5 shows an example of groundplane remapping and height reconstruction using belief propagation. The scene is based on a rough seafloor with a mean depth of 10 m. The scene also contains a large 'crater'.

V. Conclusion

Through application of the Belief Propagation algorithm to SAS imagery a three-dimensional topography of the seafloor can be obtained. This algorithm searches through a likelihood volume locating the most likely surface given a priori information regarding the likelihood of a height change between neighbouring pixels. This surface at each pixel is limited in height to that of the chosen remapped ground-planes. This produces a quantised height map resulting in large height steps and therefore height errors of up to one quantisation step. To reduce this error the algorithm need to be extended to allow for a continuous height variable at each pixel.

VI. Acknowledgments

Philip Barclay would like to thank the University of Canterbury for his Doctoral Scholarship. Chris Forne would like to thank the New Zealand Foundation for Research, Science Technology for his top achiever Doctoral Scholarship.

References

[1] D. W. Hawkins and P. T. Gough, "Imaging algorithms for a strip-map synthetic aperture sonar: minimizing the effects of aperture errors and aperture undersampling," *IEEE J. Oceanic Engineering*, vol. 22, no. 1, pp. 27–39, January 1997.

[2] F. Gatelli, A. M. Guarnieri, F. Parizzi, P. Pasquali, C. Prati, and F. Rocca, "The wavenumber shift in SAR interferometry," *IEEE Trans. Geosci. Remote Sensing*, vol. 32, no. 7, pp. 855–865, July 1994.

[3] L. M. H. Ulander and P. Frölind, "Ultra-wideband SAR interferometry," *IEEE Trans. Geoscience and Remote Sensing*, vol. 36, no. 5, pp. 1540–1550, September 1998.

[4] J. Homer, I. D. Longstaff, Z. She, and D. Gray, "High resolution 3-D imaging via multi-pass SAR," *IEE Proc. Radar, Sonar, and Navigation*, vol. 149, no. 1, pp. 45–50, February 2002.

[5] P. J. Barclay, M. P. Hayes, and P. T. Gough, "Multi-channel data acquisition for a free-towed synthetic aperture sonar," in *Proceedings of Electronics New Zealand Conference*, Dunedin, New Zealand, November 2002, pp. 133–138.

[6] M. P. Hayes, P. J. Barclay, and T. J. Hawkins, "An embedded compact PCI computer system for a synthetic aperture sonar towfish," in *Proceedings of Electronics New Zealand Conference*, Dunedin, New Zealand, November 2002, pp. 139–144.

[7] S. M. Banks, H. D. Griffiths, and T. J. Sutton, "A technique for interferometric synthetic aperture sonar image processing," in *London Communications Symposium*, September 2001.

[8] P. J. Barclay, M. P. Hayes, and P. T. Gough, "Using a multi-frequency synthetic aperture sonar for bathymetry," in *Image and Vision Computing New Zealand 2001*, University of Otago, Dunedin, New Zealand, November 2001, pp. 63–68.

[9] W. W. J. Bonifant, M. A. Richards, and J. H. McClellan, "Interferometric height estimation of the seafloor via synthetic aperture sonar in the presence of motion errors," in *IEE Proceedings - Radar, Sonar and Navigation*, vol. 147, no. 2. IEE, December 2000, pp. 322–330.

[10] L. Wang, A. Bellettini, R. Hollett, A. Tesei, and M. Pinto, "InSAS'00: Interferometric SAS and INS aided SAS imaging," in *Oceans 2001*, vol. 1. IEEE, November 2001, pp. 179–187.

[11] J. C. Dainty, Ed., *Laser Speckle and Related Phenomena*, ser. Topics in Applied Physics. Springer-Verlag, 1975, vol. 9.

[12] R. M. Goldstein, H. A. Zebker, and C. L. Werner, "Satellite radar interferometry: Two-dimensional phase unwrapping," *Radio Science*, vol. 23, no. 4, pp. 713–720, July-August 1988.

[13] D. Massonnet, H. Vadon, and M. Rossi, "Reduction of the need for phase unwrapping in radar interferometry," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, no. 2, pp. 489–497, March 1996.

[14] A. J. Hunter, M. P. Hayes, and P. T. Gough, "Simulation of wideband interferometric SAS imagery," in *Oceans 2003*, September 2003.

[15] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*, 2nd ed. Morgan Kaufmann Publishers, San Mateo, Calif., 1988.

[16] J. Sun, H. Y. Shum, and N. N. Zheng, "Stereo matching using belief propagation," in *Proceedings of the European Conference on Computer Vision*, vol. 2, Copenhagen, Denmark, May 2002, pp. 510–524.

[17] N. Petrovic, I. Cohen, B. J. Frey, R. Koetter, and T. S. Huang, "Enforcing integrability for surface reconstruction algorithms using belief propagation in graphical models," in *Computer Vision and Pattern Recognition, 2001*, vol. 1. IEEE, 2001, pp. I–743 – I–748.

[18] A. Minagawa, K. Uda, and N. Tagawa, "Region bxtraction based on belief propagation for gaussian model," in *Pattern Recognition, 2002 Proceedings*, vol. 2. IEEE, 2002, pp. 507–510.

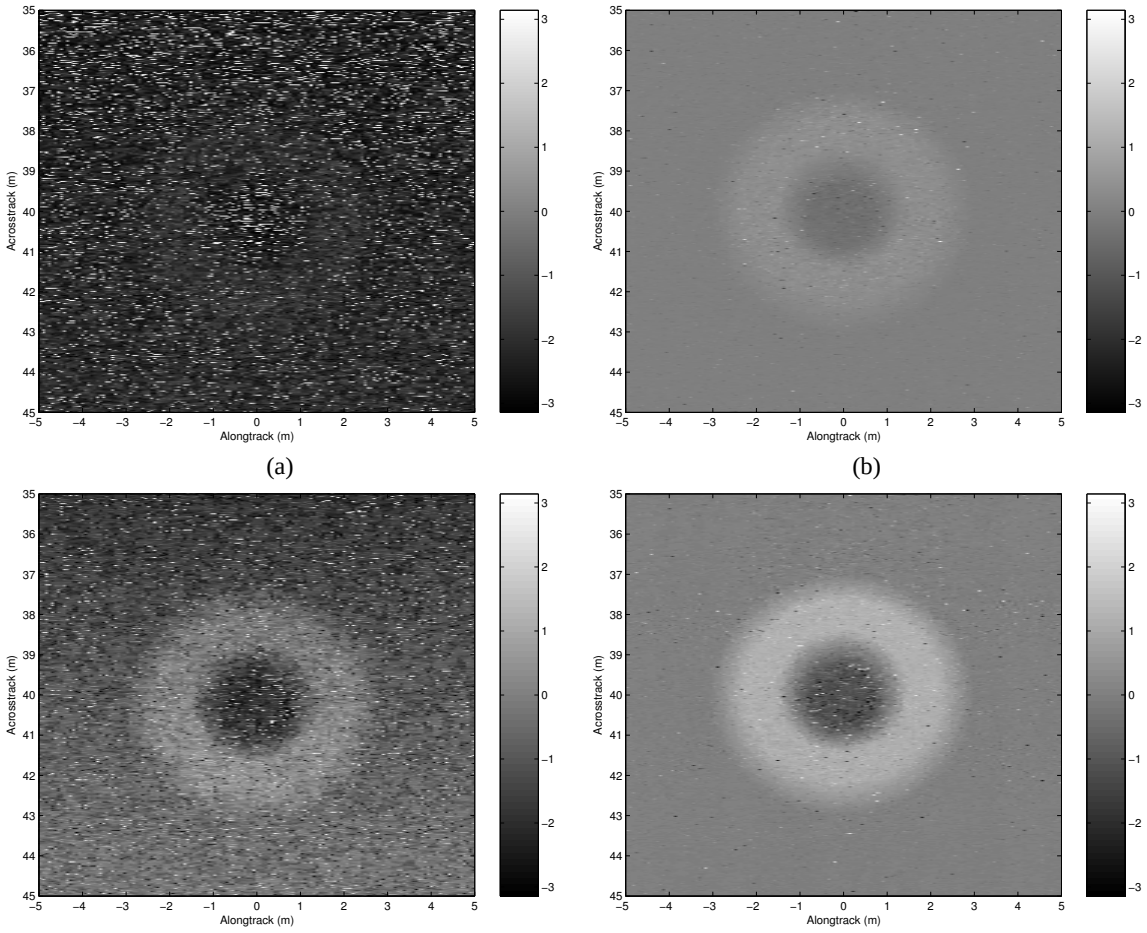


Fig. 4. Interferograms of a smaller crater of height 0.75 m on a seafloor of 10 m mean depth. (a) Slant range plane interferogram ($f_c = 30\,kHz$). (b) Remapped to estimated groundplane interferogram ($f_c = 30\,kHz$). (c) Slant range plane interferogram ($f_c = 100\,kHz$). (d) Remapped to estimated groundplane interferogram ($f_c = 100\,kHz$).

[19] Y. Weiss and W. T. Freeman, "On the optimality of solutions of the max-product belief-propagation algorithm in arbitrary graphs," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 735–744, February 2001.

[20] J. A. Goff and T. H. Jordan, "Stochastic modeling of seafloor morphology: Inversion of sea beam data for second-order statistics," *Journal of Geophysical Research*, vol. 93, no. B11, pp. 13 589–13 608, November 1988.

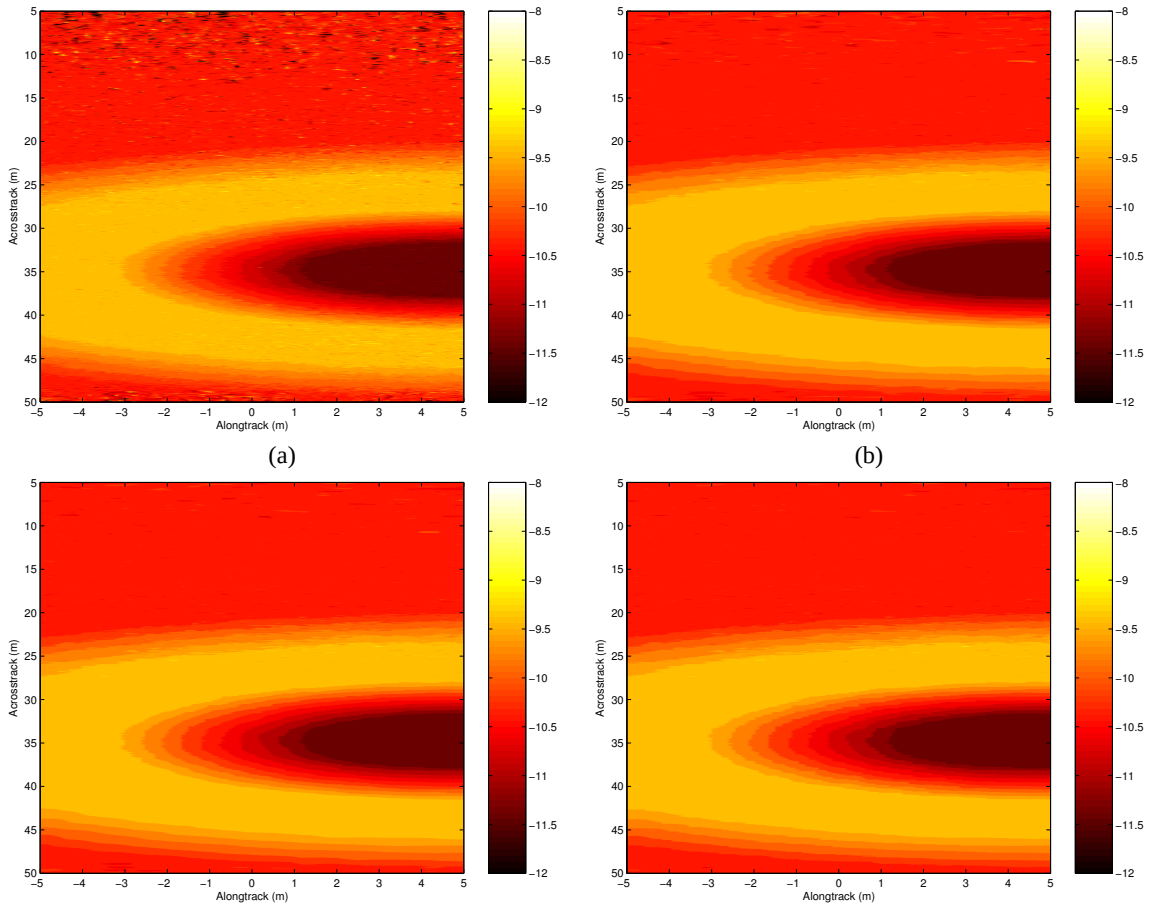


Fig. 5. Height reconstruction using Belief Propagation for a larger 'crater' of height 1 m with center −1 m on a rough seafloor of mean depth 10 m. (a) 1 iteration. (b) 5 iterations. (c) 10 iterations. (d) 20 iterations.

Appendix B

KIWISAS HARDWARE DEVELOPMENT

The KiwiSAS-II system was originally developed as a single receiver SAS. The transmitter consisted of 36 individual Tonpilz elements [Hawkins and Gough, 1996], driven as one element. The receiver consisted of a 3×3 array of hydrophones, again wired in parallel as one element. The system operated with a single linear-FM chirp of 20 kHz bandwidth, centred at 30 kHz. The signal from the hydrophone was passed along the tow-cable in analogue form to the towboat where it was down-converted to baseband, and stored for later synthetic aperture processing. After successful sea-trials with the KiwiSAS-II system, it was desired to expand the system to provide interferometric capability (KiwiSAS-III). The hydrophone array was to be rewired to three vertically separated arrays, each consisting of three parallel connected tiles. Further testing of the transmitter elements also showed a second usable frequency range in the 100 kHz range. This allowed the system to operate with two simultaneously frequency bands, one centred at 30 kHz, the second at 100 kHz. Each of these frequency bands utilised 20 kHz of bandwidth, providing equal theoretical across-track resolution from each band. The KiwiSAS-III system has been used to collect sonar data on several occasions in various locations. This data has allowed algorithms to be developed for a number of image processing stages, including synthetic aperture reconstruction, direct auto-focus [Callow et al., 2001, Callow, 2003], statistical auto-focus [Fortune et al., 2001, Fortune, 2005] and bathymetry [Barclay et al., 2001, Hayes et al., 2001]. The hardware was shown to be reliable under a range of trying conditions. More recently the system has been modified again, now recording all nine receiver elements individually [Hayes et al., 2002]. The transmitter sub-system has also been reconfigured, allowing for beam-steering of the transmitted acoustic signal [Hayes, 2003]. This system has been dubbed 'KiwiSAS-IV'. The design and construction of the interferometric capable KiwiSAS-III system is presented here, followed by a brief overview of the KiwiSAS-IV system. The description of the KiwiSAS-III system is based largely on a paper presented at the Electronics New Zealand Conference (ENZCON), 2002 [Barclay et al., 2002a].

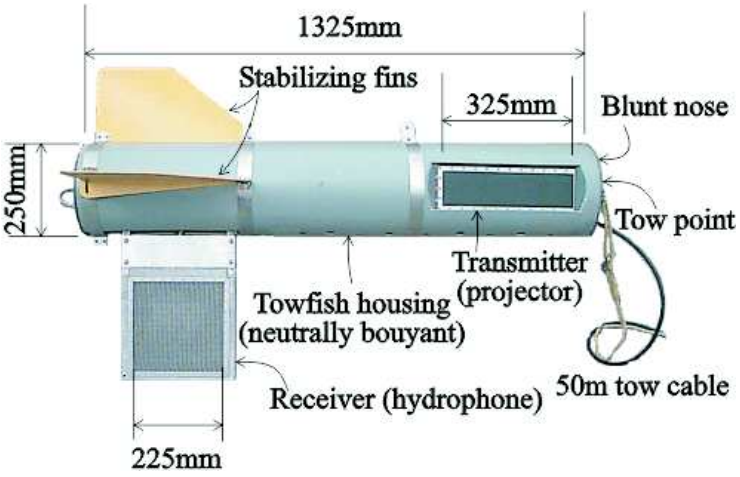


Figure B.1 KiwiSAS-III towfish.

B.1 KIWISAS-III

The KiwiSAS-III is a free-towed, side scanning sonar system developed by the Acoustics Research Group at the University of Canterbury. The system is designed as a coherent system allowing post-processing signal processing to be applied to the data. In particular, research is being performed on synthetic aperture sonar (SAS) algorithms and bathymetric terrain mapping. There are two major components to the system; the free-towed towfish, and the towboat hardware. The towfish is a nose towed, blunt nose design with stabilizing fins at the rear (see Figure B.1). These features are intended to make the towfish more stable when towed, reducing blurring of the resulting imagery. The towfish is connected to the towboat via a 50 m multi-core tow-cable. To minimise the complexity and cost of the water sealed towfish, the data acquisition hardware is located on the towboat, with all data signals to and from the towfish in an analogue form. The system is designed to operate in a continuous data acquisition mode and due to the coherent nature of the data processing, all real-time parts of the system must operate synchronously. The data acquisition hardware is mounted in a standard 19" Eurocard rack, connected to a standard PC via a custom PCI card.

B.2 TOWFISH ELECTRONICS

The towfish can be considered to have three separate sub-systems. These consist of the transmitter, receiver, and instrumentation as shown in Figure B.2. Each of these three sub-systems are independent except for a common power supply. The transmitter consists of 12 power amplifiers, each amplifier driving three Tonpilz [Hawkins and Gough, 1996] projectors. All amplifiers are driven with the same signal, so no electrical beam steering is available. The power amplifiers are powered directly off the ± 12 V

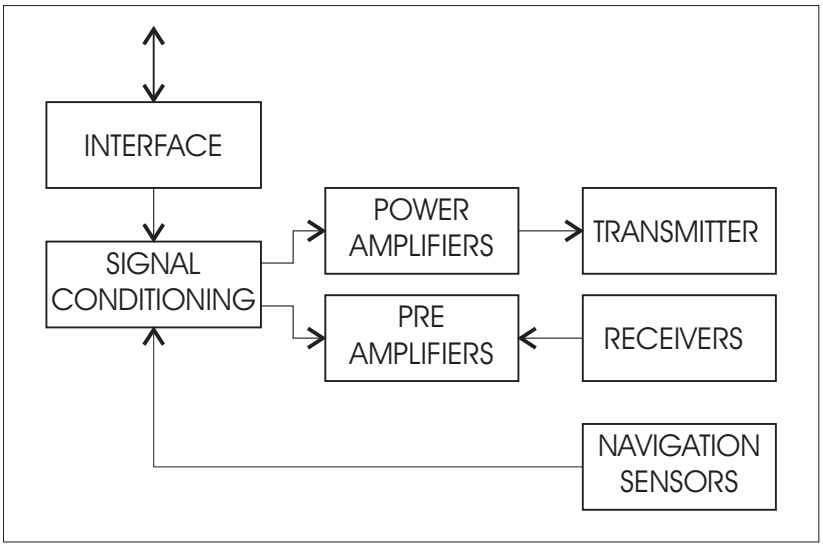


Figure B.2 Block diagram of towfish electronics.

lines of the tow-cable. Power decoupling is provided via capacitor banks of 42000 uF for each power-supply rail, located within the towfish. These capacitor banks ensure the supply voltage does not droop during the period of transmission.

The receiver hydrophone array consists of nine 75×75 mm PVDF tiles arranged in a 3×3 grid. Each row of three are wired in parallel and connected to a 40 dB pre-amplifier. The three signals are transmitted to the towboat via the tow-cable using differential line transmitters (EL2140C).

In order to estimate the orientation of the towfish an instrumentation package is included in the rear can of the towfish. This device contains an inclinometer, magnetometers, and accelerometers. These sensors are interfaced using an 8-bit microcontroller (PIC18F876A), and the resulting digital data stream sent to the towboat using NMEA type sentences via an RS-232 link [Pilbrow et al., 2002b].

B.3 TOWBOAT ELECTRONICS

Since the KiwiSAS-III sonar system is designed to be operated from a variety of small towboat vessels, the system must be simple, self contained and easy to transport. On board the boat, all equipment is housed in small, sturdy containers with large transportation handles. The overall towboat electronics setup is outlined in Figure B.3.

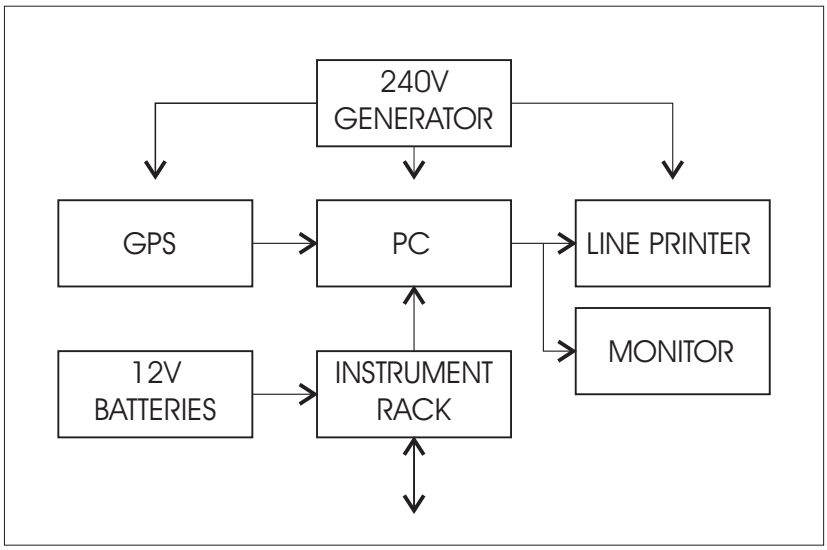


Figure B.3 Block diagram of towboat electronics.

B.3.1 Central Computer

The central computer is an off-the-shelf, dual-processor Intel® Pentium®-III system running the Linux® operating system. This computer is mounted in a sturdy metal box with large handles to ease transportation. The central computer primarily records the incoming data from the instrument rack of Section B.3.3 via a custom built PCI interface card [Hayes et al., 2002]. This data is recorded onto a standard IDE hard-disk in raw form to allow post processing to be performed.

During data collection, some form of feedback to the boat crew is required to ensure correct operation of the system. To do this, the data is pulse-compressed (correlated with the transmitted signal) and displayed on a flat screen monitor in real time. This display stream is user configurable, providing valuable information to the boat crew. The data is also recorded in a hard form using a thermal line printer. This provides a longer term history of the sonar operation. The software application running on the computer consists of two main programs. A background task provides the real time data acquisition and storage of the incoming data streams. The second task provides the visual display of this recorded data, and also any real-time processing as required. Only the raw data stream is recorded to the disk, all other processing is only used for real-time visual display.

B.3.2 Auxiliary Systems

A standard 12 channel GPS receiver on the boat provides a basic navigation recording. This positioning information is transfered to the central computer as NMEA type sentences via a serial RS-232 connection. These sentences are tagged on the central

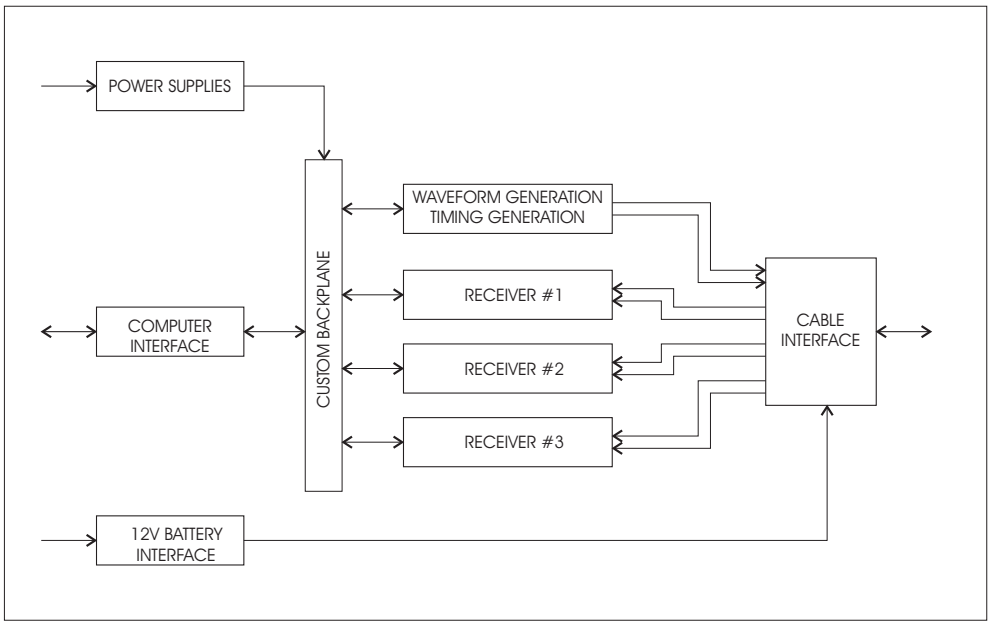


Figure B.4 Block diagram of towboat instrumentation rack electronics.

computer to synchronise with the data streams from the towfish, and recorded onto the disk. This positioning information is of limited use for the development of synthetic aperture algorithms, since the GPS receiver must be located on the towboat and not the towfish.

Power for all the systems on the boat is generated by a portable petrol 240 V AC generator. The towfish is powered separately from two large capacity, deep cycle 12 V lead-acid batteries. These batteries help reduce electrical switching noise in the towfish electronics by separating the power supplies of the towfish and the towboat.

B.3.3 Instrument Rack

The rack of electronics on-board the boat provides the heart of the real-time sonar system. This rack provides all the interfacing between the computer and the tow-cable, as outlined in Figure B.4. It is housed in a 19" Eurocard rack, utilising a custom backplane for communication between the cards. Photographs of the complete instrument rack is shown in Figure B.5. Each card in the rack is allocated a unique four-bit card number by jumpers on each card. This then allows the cards to be individually addressed from the central computer.

Backplane

The backplane is the only communications channel between the various cards within the rack. Connections from the backplane to each card are provided via 64-pin DIN

Figure B.5 Electronics rack based on a 19" Eurocard cage. The internal modules from left to right are: power supply and lead-acid battery connections, debugging port, sonar tow-cable connections and towfish power-supply current meters, three dual-channel receivers, programmable waveform transmitter, computer interface, and cooling fan. Behind the backplane is the low-voltage power supplies for the electronics.

connectors, spaced at a distance of 4 HP (1 HP=5.08 mm). The backplane is entirely passive, with three main groups of signals. The top 32 pins of each connector are connected in a parallel bus, used for digital communication between the cards. The next 14 pins are used for DC power supplies to each of the cards as outlined in Section B.3.3. The lower 18 pins of each connector are divided into six groups of three, providing analog connections to each card. Rather than these analogue pin sets being bussed along the entire backplane, connections are made to connectors on the back of the backplane. This then allows a point-to-point connection of these analogue signals to be made as required.

Power Supplies

The entire electronics rack is designed to run from a single 240 V AC power source. From this, several raw DC power supplies are derived using a single transformer with multiple secondary windings. These raw supplies are ±8 V for analog electronics supplies, +8 V for digital supplies, and ±15 V for auxiliary supplies. Each of these raw supplies are routed along the backplane to all cards, each of which linearly regulate the supplies as required. Typically, the 8 V lines are regulated to 5 V, and the 15 V lines are regulated to 12 V. By separating the analog and digital supplies at the power transformer noise is minimised. Each of the supplies has separate ground returns, and are completely isolated from each other. In order to ensure common ground potential between the analog and digital supplies, these grounds are connected by 0 Ω resistors under the analogue to digital converter (ADC) on each of the receiver cards.

Computer Interface

This is a simple card (see Figure B.6), providing buffering and termination of the high speed serial link between the computer and the backplane. It also contains a phase locked loop (PLL) chip (MC88915) to provide the main 40 MHz clock for the entire system. Buffering between the computer and backplane is achieved using HCT family buffers.

Timing/Waveform Generation Card

This card has two primary functions. The first is to generate all the timing signals to control the other parts of the system. The second is to generate the arbitrary analogue waveforms used as the acoustical transmitted signal. Central to this card is an FPGA (XILINX XC4003E), booted from an on-board serial PROM chip (XC17108E) (see Figure B.7). This FPGA device is configured with several internal registers addressable via the backplane.

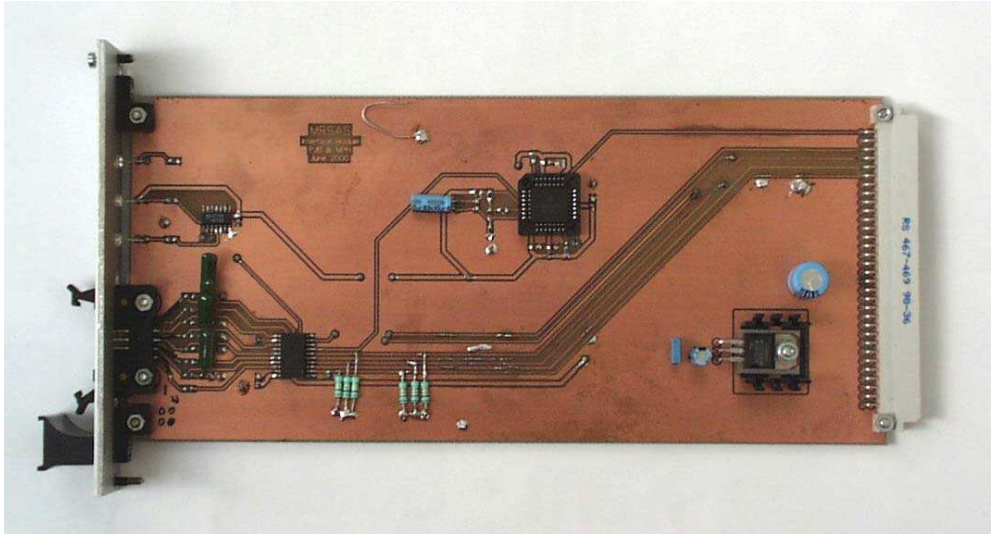


Figure B.6 Interface card from instrumentation rack to central computer.

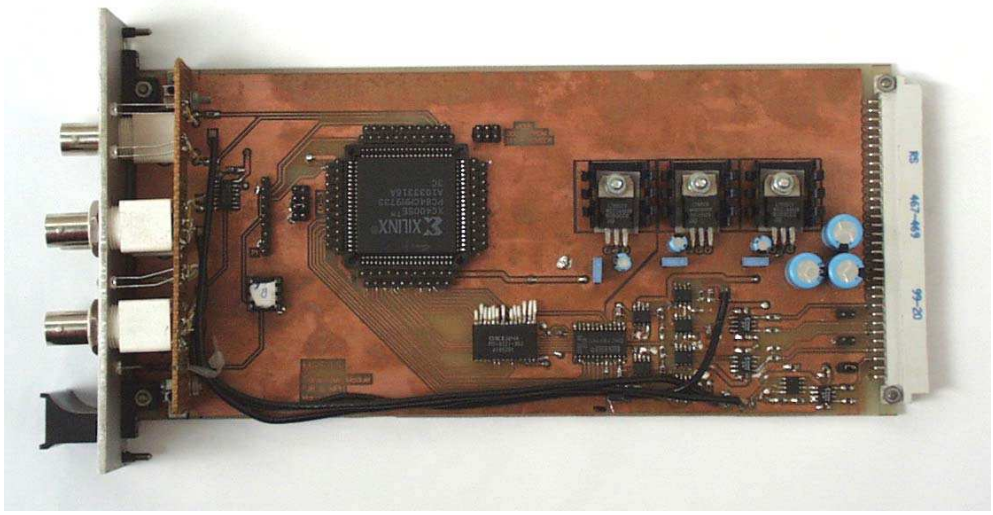


Figure B.7 Programmable transmitter module based on a XILINX XC4003E FPGA.

The overall timing of the system is controlled via two internal registers of the FPGA. These registers control the length of the transmitted waveform and the overall repetition period of the waveform. These timing signals are transmitted to the other cards in the system via the digital portion of the backplane. They are also buffered and routed to BNC test connectors on the front panel.

The arbitrary acoustic transmission signal used by the sonar are downloaded into a field memory device from the central computer, via the FPGA. These waveforms are then converted to an analog signal using a dual-channel, 12-bit, digital to analog converter (DAC) (DAC7801).

The outputs of the DAC are then buffered and passed out to the backplane via differential transmitters (EL2140C). The single ended signal is also routed to a BNC test point on the front panel, via a dedicated buffer. This allows an oscilloscope to be connected to monitor the transmitted signal. There is no gain control over any of the transmitted signals, other than varying the maximum digital code stored in the field memory. This is because the transmitter of the sonar is always operated at maximum signal levels, maximising acoustic response.

Receiver Card

The primary role of the receiver cards is to take the incoming differential signals from the towfish, and convert them into a suitable digital format for storage on the computer hard-disk. Each card has two independent, identical, receiver channels. To provide synchronicity between all the receiver channels, all devices are controlled from the same master clock and timing signals from the waveform card. Each receiver channel consists of an analog shaping stage, the digital conversion, and then digital down conversion to lower the data rate. An example of one of the three card manufactured is shown in Figure B.8.

Each channel is initially sampled as 12-bits 40 MSps, and then down-converted to 16-bit real, 16-bit imaginary samples at approximately 30 kSps. The data is sampled at a much higher than Nyquist rate and down-sampled to a lower rate, allowing low order anti-aliasing filters to be used. This simplifies the filter design, reducing the phase effects within the resulting signal spectrum.

The incoming differential signal from the towfish are terminated and converted to a single ended signal using a differential receiver (EL2142C); the matched receiver to the transmitter in the towfish. The input of the receiver may also be switched via a relay to an auxiliary test input from the backplane. This allows a known test signal to be used as the input for any of the receivers to test their operation. This relay is under software control from the PC. The single ended signal then passed though a gain stage (AD603), variable from -10 dB to +30 dB, again under software control. This allows the maximum signal to be applied to the ADC without overloading it. From

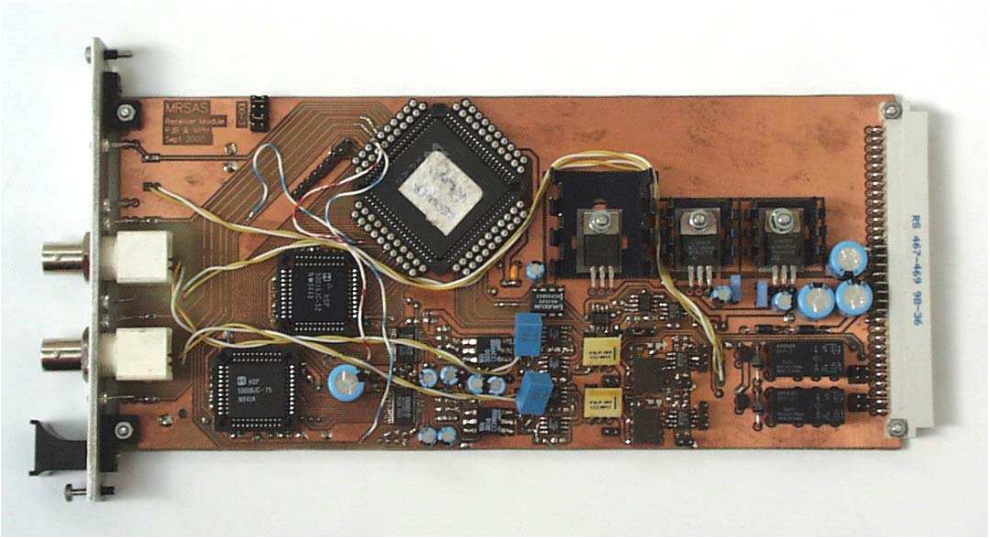


Figure B.8 Dual-channel receiver card based on a XILINX XC95108 CPLD and a pair of HSP50036 digital down-converter chips.

the output of the variable gain stage the signal is low pass filtered using a LCL filter module (P3LP-300) with a cutoff frequency of 300 kHz. This filtered signal is then applied to the input of the ADC (CLC952) via a centre-tapped transformer, driven by a unity gain buffer. The transformer is necessary to level shift the signal from a bipolar signal to a unipolar signal suitable for the input of the ADC. The ADC input is also routed to the front panel for monitoring via a dedicated buffer.

The AD converters are operating at the full 40 MHz clock frequency, performing 12-bit signal conversions. This digital data is bussed to a digital down converter (DDC) (HSP50016) chip performing a frequency shift and decimation of the data stream. The operation of the DDC is under complete software control, allowing tuning of the internal circuitry to the desired spectrum portion of interest.

This card is digitally controlled from a Complex Programmable Logic Device (CPLD) (XC95108). This device also time division multiplexes (TDM) the serial data streams from the two DDC devices. Each DDC produces a 33-bit data word, consisting of a frame pulse followed by 32 data bits. The data streams from each of the three receiver cards are then combined into one serial TDM stream, transmitted to the PC via the custom PCI card, as shown in Figure B.9.

Data from the three receiver cards is time division multiplexed together to form one data stream for transmission to the PC. This is achieved by allocating each card a unique slot number within the data stream. Each card then holds the serial lines in a tri-state condition until their time slot is reached. The card then drives the serial lines with the data from the two DDCs, then returns to a tri-state condition to allow another card to take control of the bus. This time slot allocation is achieved by setting

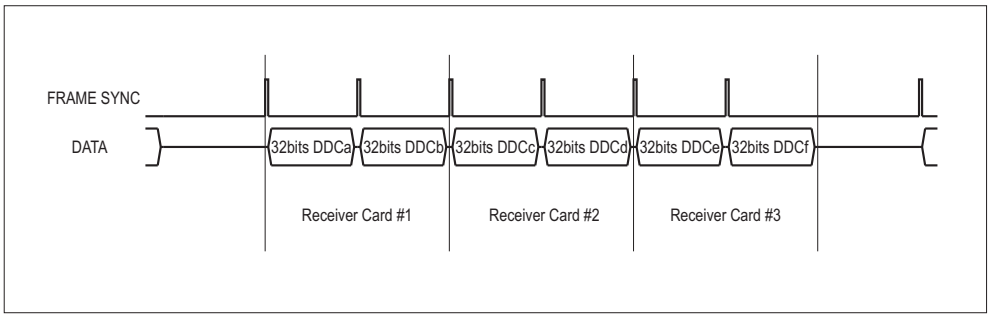


Figure B.9 Time division multiplexed bus for data.

a register within the receiver card CPLD.

The power supplies for the card are derived from the raw DC supplies of the backplane, linearly regulated to ± 5V for the analog circuits and $+5$ V for the digital circuits. This separation of the power supplies is to reduce the interference from the digital switching into the analog signals.

Digital Down Converter Operation

The main operation of the Digital Down Converter (DDC) is to perform a frequency transformation on the data stream from the receivers. This is performed to allow a more compact, efficient storage of the data set. This efficiency is gained by converting the two 20 kHz information bands of interest to baseband, drastically reducing the bandwidth of signal to be stored.

The frequency transformation is achieved by multiplying the incoming data stream by a sinusoidal wave with centre frequency equal to the centre of the frequency band of interest. The result is a frequency transformation with copies of the original spectrum centred at DC and $2 \times f_c$. By then low-pass filtering the spectrum, only the lower copy of the spectrum is required.

Control of the DDC is achieved via eight internal 40-bit registers. Each of these registers is further divided into several fields giving complete control over the device.

B.4 EXAMPLE DATA

Results from a test site in Sydney Harbour, Australia are shown in Figure B.10. This scene contains several man-made objects including several point reflectors, and a ladder-like structure. Also present is a natural bland background.

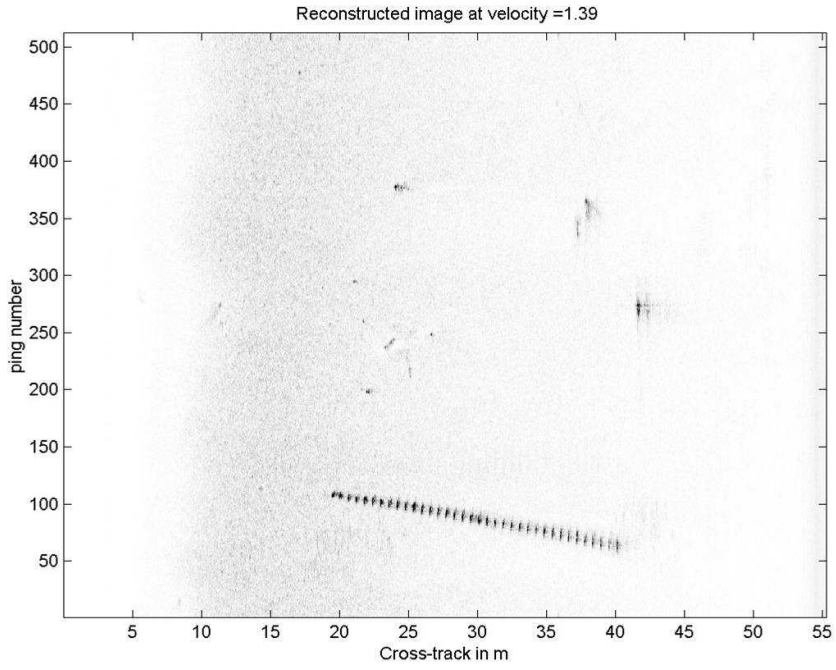


Figure B.10 Sonar image from a test site in Sydney Harbour, Australia after synthetic aperture with auto-focus processing has been applied.

B.5 KIWISAS-IV

After successful trials with the interferometric capable KiwiSAS-III system, the towfish has been redesigned to allow greater flexibility of transmitted and received signals. The transmitter subsystem has been redesigned to allow the individual columns of elements to be driven with independent signals. This will allow the transmitted beam to be steered in azimuth. The 3×3 receiver array has been rewired into individual elements, and a second array of 1×9 elements is under development; mounted along the length of the towfish on the underside. This long array will allow further work in the auto-focus field to be performed, and also allowing faster towboat operation or in increase in imaging range. The total number of receiver channels has thus been increased to a total of 18. The increase in both transmitter and receiver channels has necessitated the inclusion of an computer system within the towfish housing, and the development of a new multi-channel data acquisition system [Hayes et al., 2002]. This new system will also incorporate a new navigation system to precisely position the towfish while imaging. This system will use fixed transponders placed on the seafloor in a long baseline configuration [Pilbrow et al., 2002c, Pilbrow et al., 2002a, Pilbrow, 2007]. Photos of the new towfish housing are shown in Figure B.11, internal hardware shown in Figure B.12.

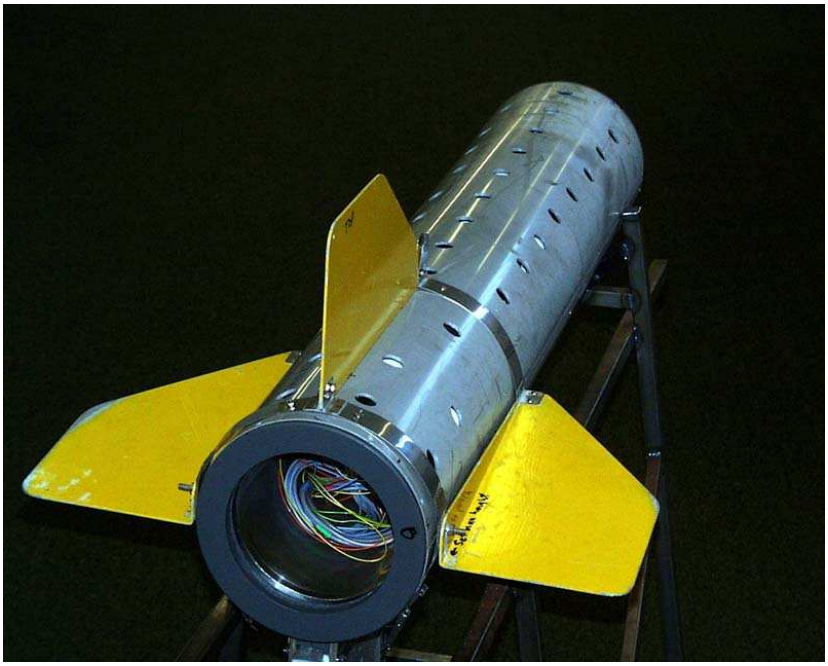


Figure B.11 KiwiSAS-IV towfish sitting on storage cradle.

B.6 COMPONENT APPENDIX

- XILINX `www.xilinx.com`
  XC95108 In-System Programmable CPLD
  XC4003 Field Programmable Gate Array
  XC17108 Serial Configuration PROM

- Analog Devices `www.analog.com`
  CLC952 12-bit, 41 MSPS Monolithic A/D Converter
  AD603 Low Noise, 90 MHz Variable-Gain Amplifier

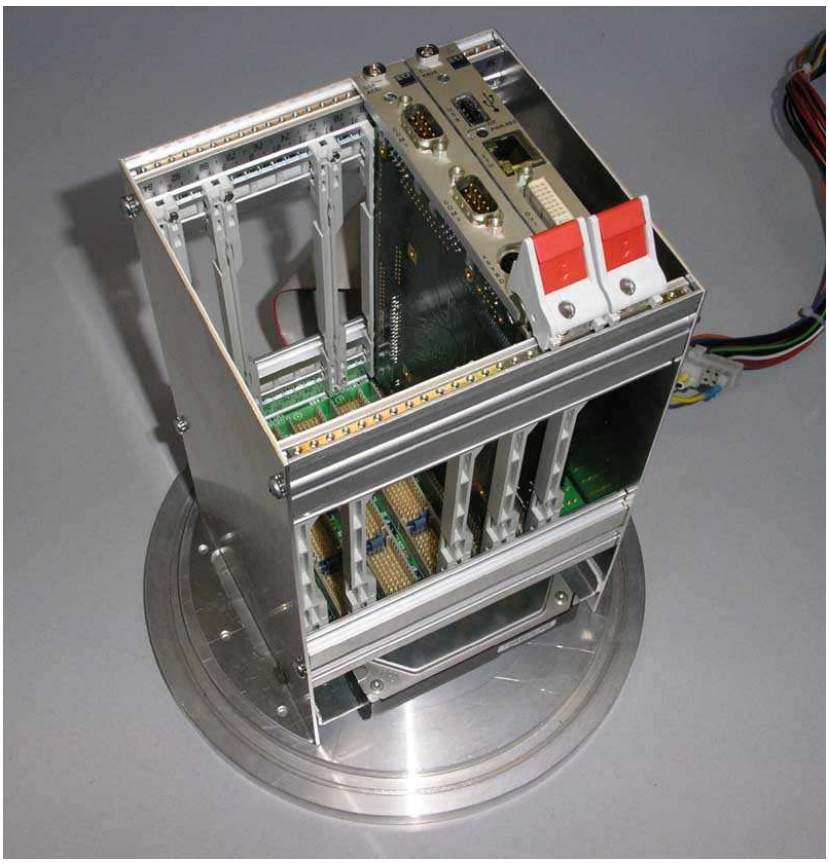
- Burr-Brown (Texas Instruments) `analog.ti.com`
  DAC7801 Dual Monolithic CMOS 12-bit Multiplying Digital to Analog Converters

- OKI Semiconductor `www.okisemi.com`
  MSM518221 282,214-Word × 8-bit Field Memory

- Motorola `www.motorola.com`
  MC88915 Low Skew CMOS PLL Clock Driver

- Harris (Intersil) `www.intersil.com`
  HSP50016 Digital Down Converter

(a) Power amplifiers.



(b) Compact-PCI cage.

Figure B.12 KiwiSAS-IV hardware internals. (a) power amplifier stack, (b) compact-PCI cage mounted atop 3.5" hard-disk. Embedded computer card loaded into cage.

- Elantec (Intersil) `www.intersil.com`
  EL2140 Differential Twisted Pair Driver
  EL2142 Differential Line Receiver

- Coilcraft `www.coilcraft.com`
  P3LP Low Pass LC Filter Modules
  WB2010 Wide Band RF Transformers

- Microchip `www.microchip.com`
  PIC16F876A 8-bit PIC microcontroller

- Intel `www.intel.com`
  Pentium-III 800 MHz mobile CPU.

REFERENCES

[Appledorn, 1996] Appledorn, C. R. (1996). A new approach to the interpolation of sampled data. *IEEE Trans. Med. Imag.*, 15(3):369–376.

[Bamler, 1992] Bamler, R. (1992). A comparison of range-Doppler and wavenumber domain SAR focusing algorithms. *IEEE Trans. Geosci. Remote Sensing*, 30(4):706–713.

[Bamler and Hartl, 1998] Bamler, R. and Hartl, P. (1998). Topical review: Synthetic aperture radar interferometry. *Inverse Problems*, 14(3):R1–R54.

[Bamler and Just, 1993] Bamler, R. and Just, D. (1993). Phase statstics and decorrelation in SAR interferograms. In *Proceeding of the Geoscience and Remote Sensing Symposium, IGARSS*, volume 3, pages 980–984.

[Banks and Griffiths, 2002] Banks, M. and Griffiths, H. D. (2002). The use of fast factorised back projection for synthetic aperture sonar imaging. In *6'th European Conf. on Underwater Acoustics, ECUA 2002*, pages 529–534, Gdansk, Poland.

[Banks, 2002] Banks, S. (2002). *Studies in High Resolution Synthetic Aperture Sonar*. PhD thesis, Department of Electrical and Electronic Engineering, University College London.

[Banks et al., 2000] Banks, S., Sutton, T., and Griffiths, H. (2000). Noise susceptibility of phase unwrapping algorithms for interferometric synthetic aperture sonar. In *Proceedings of the Fifth European Conference on Underwater Acoustics*, volume 1, pages 451–456.

[Barber, 1985] Barber, B. C. (1985). Theory of digital imaging from orbital synthetic aperture data. *International Journal of Remote Sensing*, 6(7):1009–1057.

[Barclay et al., 2003a] Barclay, P. J., Forne, C. J., Hayes, M. P., and Gough, P. T. (2003a). Reconstructing seafloor bathymetry with a multi-frequency, multi-channel broadband InSAS using belief propagation. In *Oceans 2003*. IEEE, MTS.

[Barclay et al., 2001] Barclay, P. J., Hayes, M. P., and Gough, P. T. (2001). Using a multi-frequency synthetic aperture sonar for bathymetry. In *Image and Vision*

Computing New Zealand 2001, pages 63–68, University of Otago, Dunedin, New Zealand.

[Barclay et al., 2002a] Barclay, P. J., Hayes, M. P., and Gough, P. T. (2002a). Multichannel data acquisition for a free-towed synthetic aperture sonar. In *Proceedings of Electronics New Zealand Conference*, pages 133–138, Dunedin, New Zealand.

[Barclay et al., 2002b] Barclay, P. J., Hayes, M. P., and Gough, P. T. (2002b). Reconstructing seafloor bathymetry with a multichannel broadband InSAS. In *Image and Vision Computing New Zealand 2002*, pages 47–52, University of Auckland, Auckland, New Zealand.

[Barclay et al., 2003b] Barclay, P. J., Hayes, M. P., and Gough, P. T. (2003b). Bathymetry reconstruction for a free-towed synthetic aperture sonar. In *Proceedings of the World Congress on Ultrasonics*, pages 519–522.

[Barclay et al., 2004] Barclay, P. J., Hayes, M. P., and Gough, P. T. (2004). Multi-receiver, multi-frequency maximum likelihood synthetic aperture sonar interferometry. In *Proceedings of Image and Vision Computing New Zealand 2004*, pages 273–278, Akaroa, New Zealand.

[Barclay et al., 2005] Barclay, P. J., Hayes, M. P., and Gough, P. T. (2005). ML estimation of seafloor topography using multi-frequency synthetic aperture sonar. In *Oceans 2005 Europe*, Brest, France. IEEE.

[Barclay et al., 2006] Barclay, P. J., Hayes, M. P., and Gough, P. T. (2006). Bathymetric results from a multi-frequency InSAS sea-trial. In *Oceans 2006 Asia-Pacific*, Singapore. IEEE, OES.

[Bellec et al., 2005] Bellec, R., Legris, M., Khenchaf, A., Amate, M., and Hetet, A. (2005). Repeat-track SAS interferometry: feasibility study. In *Proceedings of MTS/IEEE Oceans 2005*, volume 1, pages 748–754. MTS/IEEE.

[Bellettini and Pinto, 2002] Bellettini, A. and Pinto, M. A. (2002). Theoretical accuracy of synthetic aperture sonar micronavigation using a displaced phase-center antenna. *IEEE J. Oceanic Eng.*, 27(4):780–789.

[Berizzi et al., 1997] Berizzi, F., Corsini, G., Diani, M., Lombardini, F., and Pinelli, G. (1997). Simulation model and performance analysis of a three-antenna InSAR system. In *IEE Radar 97*, number 449, pages 119–123. IEE.

[Blanc-Benon and Jauffret, 1997] Blanc-Benon, P. and Jauffret, C. (1997). TMA from bearings and multipath time delays. *IEEE Trans. Aerosp. Electron. Syst.*, 33(3):813–824.

[Bonifant, 1999] Bonifant, W. W. J. (1999). Interferometric synthetic aperture sonar processing. Master's thesis, Georgia Institute of Technology.

[Bonifant et al., 1999] Bonifant, W. W. J., Richards, M. A., and McClellan, J. H. (1999). An analysis of the effect of motion and phase errors on the implementation of interferometric processing by synthetic aperture sonar. In *Signals, Systems, and Computers, 1999. Conference Record of the Thirty-Third Asilomar Conference on*, volume 1, pages 269–273.

[Born and Wolf, 1999] Born, M. and Wolf, E. (1999). *Principles of Optics*. Cambridge University Press, 6th edition.

[Cafforio et al., 1991] Cafforio, C., Prati, C., and Rocca, E. (1991). SAR data focusing using seismic migration techniques. *IEEE Trans. Aerosp. Electron. Syst.*, 27(2):194–207.

[Callow, 2003] Callow, H. J. (2003). *Signal Processing for Synthetic Aperture Sonar Image Enhancement*. PhD thesis, Department of Electrical and Electronic Engineering, University of Canterbury, New Zealand.

[Callow et al., 2001] Callow, H. J., Hayes, M. P., and Gough, P. T. (2001). Autofocus of multi-band, shallow-water, synthetic aperture sonar imagery using shear-averaging. In *IGARSS 2001*. IEEE.

[Carrara et al., 1995] Carrara, W. G., Goodman, R. S., and Majewski, R. M. (1995). *Spotlight synthetic aperture radar: signal processing algorithms*. Artech House.

[Carter et al., 1973] Carter, G. C., Knapp, C. H., and Nuttall, A. H. (1973). Estimation of the magnitude-squared coherence function via overlapped fast fourier transform processing. *IEEE Trans. Audio Electroacoust.*, 21(4):337–389.

[Chang et al., 1992] Chang, C. Y., Jin, M. Y., and Curlander, J. C. (1992). SAR processing based on the exact two-dimensional transfer function. In *Geoscience and Remote Sensing Symposium*, pages 355–359.

[Châtillon et al., 1999] Châtillon, J., Adams, A. E., Lawlor, M. A., and Zakharia, M. E. (1999). SAMI: A low-frequency prototype for mapping and imaging of the seabed by means of synthetic aperture. *IEEE J. Oceanic Eng.*, 24(1):4–15.

[Chen and Zebker, 2000] Chen, C. W. and Zebker, H. A. (2000). Two-dimensional phase unwrapping with statistical models for nonlinear optimization. In *Geoscience and Remote Sensing Symposium Proceesings, 2000, Proceedings IGARSS 2000*, volume 7, pages 3213–3215. IEEE International.

[Chen and Zebker, 2002] Chen, C. W. and Zebker, H. A. (2002). Phase unwrapping for large SAR interferograms: Statistical segmentation and generalized network models. *IEEE Trans. Geosci. Remote Sensing*, 40(8):1709–1719.

[Colesanti et al., 2003] Colesanti, C., Ferretti, A., Novali, F., Prati, C., and Rocca, F. (2003). SAR monitoring of progressive and seasonal ground deformation using the permanent scatters technique. *IEEE Trans. Geosci. Remote Sensing*, 41(7):1685–1701.

[Corsini et al., 1997] Corsini, G., Diani, M., Lombardini, F., and Pinelli, G. (1997). Reduction of the phase-unwrapping drawbacks by the three-antenna interferometric SAR system. In *Geoscience and Remote Sensing, 1997. IGARSS '97*, volume 4, pages 1536–1538. IEEE International.

[Corsini et al., 1999] Corsini, G., Diani, M., Lombardini, F., and Pinelli, G. (1999). Simulated analysis and optimization of a three-antenna airborne InSAR system for topographic mapping. *IEEE Trans. Geosci. Remote Sensing*, 37(5):2518–2529.

[Cumming et al., 1992] Cumming, I., Wong, F., and Raney, K. (1992). A SAR processing algorithm with no interpolation. In *Geoscience and Remote Sensing Symposium*, pages 376–379.

[Curlander and McDonough, 1991] Curlander, J. C. and McDonough, R. N. (1991). *Synthetic Aperture Radar: systems and signal processing*. John Wiley & Sons, Inc.

[Davenport and Root, 1987] Davenport, W. B. J. and Root, W. L. (1987). *An introduction to the theory of random signals and noise*. IEEE Press.

[Deffenbaugh et al., 1996] Deffenbaugh, M., Bellingham, J. G., and Schmidt, H. (1996). The relationship between spherical and hyperbolic positioning. In *OCEANS '96. MTS/IEEE. 'Prospects for the 21st Century'. Conference Proceedings*, volume 2, pages 590–595.

[Dix and Palmer, 1984] Dix, J. F. and Palmer, R. F. (1984). Study of the relative sonar performance of incoherent and coherent processing against echo fading in shallow water. *Proc. IEE*, 131, Part F:308–314.

[Dybdal, 1986] Dybdal, R. B. (1986). Monopulse resolution of interferometric ambiguities. *IEEE Trans. Aerosp. Electron. Syst.*, AES-22(2):177–183.

[Fortune, 2005] Fortune, S. A. (2005). *Phase error estimation for synthetic aperture imagery*. PhD thesis, Department of Electrical and Electronic Engineering, University of Canterbury, New Zealand.

[Fortune et al., 2001] Fortune, S. A., Hayes, M. P., and Gough, P. T. (2001). Statistical autofocus of synthetic aperture sonar images using image contrast optimisation. In *Oceans 2001*, volume 1, pages 163–169. IEEE.

[Fowler, 1993] Fowler, A. (1993). *The Chicken or the Egg?* Scholastic Library Publishing.

[Gabriel and Goldstein, 1988] Gabriel, A. and Goldstein, R. M. (1988). Crossed orbit interferometry: theory and experimental results from SIR-B. *Int. J. Remote Sens.*, 9(5):857–872.

[Gatelli et al., 1994] Gatelli, F., Guarnieri, A. M., Parizzi, F., Pasquali, P., Prati, C., and Rocca, F. (1994). The wavenumber shift in SAR interferometry. *IEEE Trans. Geosci. Remote Sensing*, 32(7):855–865.

[Ghiglia and Pritt, 1998] Ghiglia, D. C. and Pritt, M. D. (1998). *Two-dimensional phase unwrapping: theory, algorithms, and software.* Wiley, New York.

[Ghiglia and Wahl, 1994] Ghiglia, D. C. and Wahl, D. E. (1994). Interferometric synthetic aperture radar terrain elevation mapping from multiple observations. In *Digital Signal Processing Workshop*, pages 33–36. IEEE.

[Gini et al., 2002] Gini, F., Lombardini, F., and Montanari, M. (2002). Layover solution in multibaseline SAR interferometry. *IEEE Trans. Aerosp. Electron. Syst.*, 38(4):1344–1356.

[Goldstein et al., 1988] Goldstein, R. M., Zebker, H. A., and Werner, C. L. (1988). Satellite radar interferometry: Two-dimensional phase unwrapping. *Radio Science*, 23(4):713–720.

[Goodman, 1985] Goodman, J. W. (1985). Statistical properties of laser speckle patterns. In Dainty, J. C., editor, *Laser speckle and related phenomena*, pages 9–75. Springer-Verlag, Berlin.

[Gough and Hawkins, 1997] Gough, P. T. and Hawkins, D. W. (1997). Imaging algorithms for a strip-map synthetic aperture sonar: Minimizing the effects of aperture errors and aperture undersampling. *IEEE J. Oceanic Eng.*, 22(1):27–39.

[Gough et al., 2006] Gough, P. T., Hayes, M. P., and Noonchester, M. A. (2006). Towards the imaging of targets buried beneath the seafloor using a two-band, interferometric synthetic aperture sonar. In *Proceedings of the Eighth European Conference on Underwater Acoustics, 8th ECUA*, Carvoeiro, Portugal.

[Gough and Miller, 2003] Gough, P. T. and Miller, M. (2003). Improving the quality of synthetic aperture sonar images using data-driven autofocus techniques. In *5'th World Congress on Ultrasonics*, pages 677–680, Paris, France.

[Gough and Miller, 2004] Gough, P. T. and Miller, M. A. (2004). Displaced ping imaging autofocus for a multi-hydrophone SAS. *IEE Proc.-Radar Sonar Navig.*, 151(3):163–170.

[Graham, 1974] Graham, L. C. (1974). Synthetic interferometer radar for topographic mapping. *Proc. IEEE*, 62:763–768.

[Griffiths et al., 1997] Griffiths, H. D., Rafik, T. A., Meng, Z., Cowan, C. F. N., Shafeeu, H., and Anthony, D. K. (1997). Interferometric synthetic aperture sonar for high resolution 3-D mapping of the seabed. *IEE Proc.- Radar, Sonar and Navig.*, 144(2):96–103.

[Guarnieri and Prati, 1997] Guarnieri, A. M. and Prati, C. (1997). SAR interferometry: A "Quick and Dirty" coherence estimator for data browsing. *IEEE Trans. Geosci. Remote Sensing*, 35(3):660–669.

[Gustavsson et al., 1993] Gustavsson, A., Frölind, P. O., Hellsten, H., Jonsson, T., Larsson, B., and Stenström, G. (1993). The airborne VHF SAR system CARABAS. In *Proceedings of IGARSS'93*, volume 2, pages 558–562, Tokyo, Japan.

[Hanssen and Bamler, 1999] Hanssen, R. and Bamler, R. (1999). Evaluation of interpolation kernels for SAR interferometry. *IEEE Trans. Geoscience and Remote Sensing*, 37(1):318–321.

[Hawkins, 1996] Hawkins, D. W. (1996). *Synthetic aperture imaging algorithms: With application to wide bandwidth sonar.* PhD thesis, Department of Electrical and Electronic Engineering, University of Canterbury, New Zealand.

[Hawkins and Gough, 1995] Hawkins, D. W. and Gough, P. T. (1995). Recent sea trials of a synthetic aperture sonar. *Proceedings of the Institute of Acoustics*, 17(8):1–10.

[Hawkins and Gough, 1996] Hawkins, D. W. and Gough, P. T. (1996). Multi-resonance design of a Tonpilz transducer using the finite element method. *IEEE Trans. Ultrasonics, Ferroelectrics, and Frequency Control*, 43(5):231–239.

[Hawkins and Gough, 1997] Hawkins, D. W. and Gough, P. T. (1997). An accelerated chirp scaling algorithm for synthetic aperture imaging. In *Geoscience and Remote Sensing Symposium*, volume 1, pages 471–473.

[Hayes, 2003] Hayes, M. P. (2003). A real-time beam steering system for KiwiSAS-IV. In *Proceeding of Electronics New Zealand Conference, ENZCON2003*, pages 27–32, Hamilton, New Zealand.

[Hayes, 2004] Hayes, M. P. (2004). Multipath reduction with a three element interferometric synthetic aperture sonar. In *Proceedings of the Seventh European Conference on Underwater Acoustics*, Delft, The Netherlands.

[Hayes and Barclay, 2003] Hayes, M. P. and Barclay, P. J. (2003). The effects of multipath on a bathymetric synthetic aperture sonar using belief propagation. In *Proceedings of Image and Vision Computing New Zealand, IVCNZ2003*, Palmerston North, New Zealand.

[Hayes et al., 2001] Hayes, M. P., Barclay, P. J., Gough, P. T., and Callow, H. J. (2001). Test results from a multi-frequency bathymetric synthetic aperture sonar. In *Oceans 2001*, volume 1, pages 1682–1688. IEEE.

[Hayes et al., 2002] Hayes, M. P., Barclay, P. J., and Hawkins, T. J. (2002). An embedded compact PCI computer system for a synthetic aperture sonar towfish. In *Proceedings of Electronics New Zealand Conference*, pages 139–144, Dunedin, New Zealand.

[Hayes and Gough, 1992] Hayes, M. P. and Gough, P. T. (1992). Broad-band synthetic aperture sonar. *IEEE J. Oceanic Eng.*, 17(1):80–94.

[Hayes and Gough, 1999] Hayes, M. P. and Gough, P. T. (1999). Imaging with a multiple-hydrophone synthetic aperture sonar. In *Image Vision Computing New Zealand 1999*. IVCNZ.

[Hayes and Ho, 2000] Hayes, M. P. and Ho, T. Y. (2000). Height estimation of a sonar towfish from sidescan imagery. In *Image Vision Computing New Zealand 2000*, pages 102–107, Hamilton, New Zealand.

[Hayes et al., 2005] Hayes, M. P., Hunter, A. J., Barclay, P. J., and Gough, P. T. (2005). Estimating layover in broadband synthetic aperture sonar bathymetry. In *Oceans 2005 Europe*, Brest, France. IEEE.

[Homer et al., 2002] Homer, J., Longstaff, I. D., She, Z., and Gray, D. (2002). High resolution 3-D imaging via multi-pass SAR. *IEE Proc.-Radar Sonar Navig.*, 149(1):45–50.

[Houston, 1994] Houston, K. M. (1994). A fast beamforming algorithm. In *IEEE Oceans Conf.*, volume 1, pages 211–216.

[Hunter, 2006] Hunter, A. J. (2006). *Underwater acoustic modelling for synthetic aperture sonar*. PhD thesis, Department of Electrical and Computer Engineering, University of Canterbury, New Zealand.

[Hunter and Hayes, 2005] Hunter, A. J. and Hayes, M. P. (2005). Towards more accurate shadow modelling for simulated SAS imagery. In *Oceans 2005 Europe*, Brest, France. IEEE.

[Hunter et al., 2003] Hunter, A. J., Hayes, M. P., and Gough, P. T. (2003). Simulation of wideband interferometric SAS imagery. In *Oceans 2003*.

[Institute of Electrical and Electronic Engineers, 1985] Institute of Electrical and Electronic Engineers (1985). ANSI/IEEE standard 754-1955, IEEE Standard for binary floating-point arithmetic.

[Jenkins and Watts, 1968] Jenkins, G. M. and Watts, D. G. (1968). *Spectral analysis and its applications*. Holden Day series in time series analysis. Holden-Day, San Francisco.

[Jin and Tang, 1996] Jin, G. and Tang, D. (1996). Uncertainties of differential phase estimation associated with interferometric sonars. *IEEE J. Oceanic Engineering*, 21(1):53–63.

[Jin and Wu, 1984] Jin, M. Y. and Wu, C. (1984). A SAR correlation algorithm wich accommodates large-range migration. *IEEE Trans. Geosci. Remote Sensing*, 22(6):592–597.

[Just and Bamler, 1994] Just, D. and Bamler, R. (1994). Phase statistics of interferograms with applications to synthetic aperture radar. *Applied Optics*, 33(20):4361–4368.

[Kim and Griffiths, 1999] Kim, M. G. and Griffiths, H. D. (1999). Phase unwrapping of multibaseline interferometry using Kalman filtering. In *Image Processing and Its Applications, 1999. Seventh International Conference on*, volume 2, pages 813–817.

[Kliger and Olenberger, 1975] Kliger, I. E. and Olenberger, C. F. (1975). Multiple target effects on monopulse signal processing. *IEEE Trans. Aerosp. Electron. Syst.*, AES-11(5):795–803.

[Knab, 1983] Knab, J. J. (1983). The sampling window. *IEEE Trans. Inform. Theory*, pages 157–159.

[Knapp and Carter, 1976] Knapp, C. H. and Carter, G. C. (1976). The generalised correlation method for estiamtion of time delay. *IEEE Trans. Acoust., Speech, Signal Processing*, 24(4):320–327.

[Kraeutner and Bird, 1999] Kraeutner, P. H. and Bird, J. S. (1999). Beyond interferometry, resolving multiple angles-of-arrival in swath bathymetric imaging. In *Oceans '99 MTS/IEEE. Riding the Crest into the 21st Century*, volume 1, pages 37–45. IEEE/MTS.

[Lanari et al., 1996] Lanari, R., Fornaro, G., Riccio, D., Migliaccio, M., Papathanassiou, K. P., Moreira, J. R., Schwäbisch, M., Dutra, L., Puglisi, G., Franceschetti, G., and Coltelli, M. (1996). Generation of digital elevation models by using SIR-C/X-SAR multifrequency two-pass interferometry: The Etna case study. *IEEE Trans. Geosci. Remote Sensing*, 34(5):1097–1114.

[Lawlor et al., 1996] Lawlor, M. A., Adams, A. E., Hinton, O. R., Riyait, V. S., and Sharif, B. S. (1996). Further results from the SAMI synthetic aperture sonar. In *OCEANS '96. MTS/IEEE. 'Prospects for the 21st Century'. Conference Proceedings*, volume 2, pages 545–550.

[Lee et al., 1993] Lee, J.-S., Hoppel, K. W., Mango, S. A., and Miller, A. R. (1993). Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery. *IEEE Trans. Geosci. Remote Sensing*, 32(5):1017–1027.

[Lee et al., 2004] Lee, P. M., Jeon, B. H., Kim, S. M., Choi, H. T., Lee, C. M., Aoki, T., and Hyakudome, T. (2004). An integrated navigation system for autonomous underwater vehicles with two range sonars, inertial sensors and Doppler velocity log. In *Proceedings of Techo-Oceans*, volume 3, pages 1586–1593. MTS/IEEE.

[Li and Goldstein, 1990] Li, F. K. and Goldstein, R. M. (1990). Studies of multibaseline spaceborne interferometric synthetic aperture radars. *IEEE Trans. Geosci. Remote Sensing*, 28(1):88–97.

[Li and Stoica, 1999] Li, J. and Stoica, P. (1999). Efficient mixed-spectrum estimation with applications to target feature extraction. *IEEE Trans. Signal Processing*, 44:281–295.

[Lombardini, 1996] Lombardini, F. (1996). Absolute phase retrieval in a three-element synthetic aperture sonar interferometer. In *Radar, 1996. Proceedings. CIE International Conference of*, pages 309–312.

[Lombardini, 1998] Lombardini, F. (1998). Optimum absolute phase retrieval in three-element SAR interferometer. *IEE Electron. Lett.*, 34(15):1522–1524.

[Lombardini and Lombardo, 1996] Lombardini, F. and Lombardo, P. (1996). Maximum likelihood array SAR interferometry. In *Proc. 1996 IEEE Digital Signal Processing Workshop*, pages 358–361.

[Lombardini and Lombardo, 1998] Lombardini, F. and Lombardo, P. (1998). A ML thinned array SAR interferometric sensor for high accuracy absolute phase retrieval. In *Radar Conference, 1998. RADARCON 98 Proceedings of*, pages 263–268, 11-14 May. IEEE.

[Lombardo and Lombardini, 1997] Lombardo, P. and Lombardini, F. (1997). Multibaseline SAR interferometry for terrain slope adaptivity. In *Radar Conference, 1997 IEEE National*, pages 196–201.

[Lurton, 2000] Lurton, X. (2000). Swath bathymetry using phase difference: Theoretical analysis of acoustical measurement precision. *IEEE J. Oceanic Engineering*, 25(3):351–363.

[Maeland, 1988] Maeland, E. (1988). On the comparison of interpolation methods. *IEEE Trans. Med. Imag.*, 7(3):213–217.

[Massonnet et al., 1996] Massonnet, D., Vadon, H., and Rossi, M. (1996). Reduction of the need for phase unwrapping in radar interferometry. *IEEE Trans. Geosci. Remote Sensing*, 34(2):489–497.

[Migliaccio and Bruno, 2003] Migliaccio, M. and Bruno, F. (2003). A new interpolation kernel for SAR interferometric registration. *IEEE Trans. Geoscience and Remote Sensing*, 41(5):1105–1110.

[Milne, 1983] Milne, P. H. (1983). *Underwater acoustic positioning systems.* Spon, London.

[Minagawa et al., 2002] Minagawa, A., Uda, K., and Tagawa, N. (2002). Region extraction based on belief propagation for gaussian model. In *Pattern Recognition, 2002 Proceedings*, volume 2, pages 507–510. IEEE.

[Moreira, 1990] Moreira, A. (1990). An improved multi-look technique to produce SAR imagery. In *Record of the IEEE 1990 International Radar Conference*, pages 57–63.

[Parker et al., 1983] Parker, J. A., Kenyon, R. V., and Troxel, D. E. (1983). Comparison of interpolating methods for image resampling. *IEEE Trans. Med. Imag.*, MI-2:31–39.

[Pearl, 1988] Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: networks of plausible inference.* Morgan Kaufmann Publishers, San Mateo, Calif., 2nd revised edition.

[Petrovic et al., 2001] Petrovic, N., Cohen, I., Frey, B. J., Koetter, R., and Huang, T. S. (2001). Enforcing integrability for surface reconstruction algorithms using belief propagation in graphical models. In *Computer Vision and Pattern Recognition, 2001*, volume 1, pages I–743 – I–748. IEEE.

[Pilbrow, 2007] Pilbrow, E. N. (2007). *Synthetic aperture sonar micronavigation using an active acoustic beacon.* PhD thesis, Department of Electrical and Computer Engineering, University of Canterbury, New Zealand.

[Pilbrow et al., 2002a] Pilbrow, E. N., Gough, P. T., and Hayes, M. P. (2002a). Dual-transponder precision navigation system for synthetic aperture sonar. In *Proceedings of Electronics New Zealand Conference*, Dunedin, New Zealand.

[Pilbrow et al., 2002b] Pilbrow, E. N., Gough, P. T., and Hayes, M. P. (2002b). Inertial navigation system for a synthetic aperture sonar towfish. In *Proceedings of Electronics New Zealand Conference*, Dunedin, New Zealand.

[Pilbrow et al., 2002c] Pilbrow, E. N., Gough, P. T., and Hayes, M. P. (2002c). Long baseline precision navigation system for synthetic aperture sonar. In *Proceedings of The 11th Australasian Remote Sensing and Photogrammetry Association Conference*, Brisbane, Australia.

[Pinto et al., 2004] Pinto, M., Bellettini, A., Wang, L. S., Munk, P., Myers, V., and Pautet, L. (2004). A new synthetic aperture sonar design with multipath mitigation. In *Proceedings of the High-Frequency Ocean Acoustics Conference*, La Jolla, California.

[Prati et al., 1990] Prati, C., Rocca, F., Guarnieri, A. M., and Damonti, E. (1990). Seismic migration for SAR focusing: Interferometrical applications. *IEEE Trans. Geosci. Remote Sensing*, 28(4):627–640.

[Raney et al., 1994] Raney, R. K., Runge, H., Bamler, R., Cumming, I. G., and Wong, F. H. (1994). Precision SAR processing using chirp scaling. *IEEE Trans. Geosci. Remote Sensing*, 32(4):786–799.

[Reed et al., 2003] Reed, S., Petillot, Y., and Bell, J. (2003). An automatic approach to the detection and extraction of mine features in sidescan sonar. *IEEE J. Oceanic Eng.*, 28(1):90–105.

[Reed et al., 2004] Reed, S., Petillot, Y., and Bell, J. (2004). Automated approach to classification of mine-like objects in sidescan sonar using highlight and shadow information. *IEE Proc.-Radar Sonar Navig.*, 151(1):48–56.

[Reson, 2006] Reson (2006). SeaBat 7125 product specification, high-resolution multibeam echosounder system. Technical report, Reson, Denmark. http://www.reson.com/graphics/design/Spec sheets/SeaBat/7125.pdf.

[Rodriguez and Martin, 1992] Rodriguez, E. and Martin, J. M. (1992). Theory and design of interferometric synthetic aperture radars. *Radar and Signal Processing, IEE Proceedings of*, 139(2):147–159.

[Rolt and Schmidt, 1992] Rolt, K. D. and Schmidt, H. (1992). Azimuthal ambiguities in synthetic aperture sonar and synthetic aperture radar imagery. *IEEE J. Oceanic Eng.*, 17(1):73–79.

[Rosen et al., 2000] Rosen, P. A., Hensley, S., Joughin, I. R., Li, F. K., Madsen, S. N., Rodríguez, E., and Goldstein, R. M. (2000). Synthetic aperture radar interferometry. *Proc. IEEE*, 88(3):333–382.

[Runge and Bamler, 1992] Runge, H. and Bamler, R. (1992). A novel high precision SAR focussing algorithm based on chirp scaling. In *Geoscience and Remote Sensing Symposium*, pages 372–375.

[Scheiber and Bothale, 2002] Scheiber, R. and Bothale, V. M. (2002). Interferometric multi-look techniques for SAR data. In *Proceedings of the Geoscience and Remote Sensing Symposium, IGARSS*, volume 1, pages 173–175. IEEE International.

[Schock et al., 2005] Schock, S. G., Wulf, J., Quentina, G., and Sara, J. (2005). Synthetic aperture processing of buried object scanning sonar data. In *Proceedings of the MTS/IEEE Oceans Conference*, volume 3, pages 2236–2241. MTS/IEEE.

[Seymour and Cumming, 1994] Seymour, M. S. and Cumming, L. G. (1994). Maximum likelihood estimation for SAR interferometry. In *Geoscience and Remore Sensing Symposium, IGARSS '94*, volume 4, pages 2272–2275.

[Shippey et al., 2005] Shippey, G., Banks, S., and Pihl, J. (2005). SAS image reconstruction using fast polar back projection: Comparisons with fast factored back projection and Fourier-domain imaging. In *Proceedings of Oceans Europe*, pages 96–101, Brest, France. IEEE.

[Shippey et al., 1998] Shippey, G. A., Ulriksen, P., and Liu, Q. (1998). Quasi-narrowband processing of wideband sonar echoes. In *Proceedings of the 4 th European Conference on Underwater Acoustics (ECUA)*, pages 63–68, Rome.

[Soumekh, 1992] Soumekh, M. (1992). A system model and inversion for synthetic aperture radar imaging. *IEEE Trans. Image Processing*, 1(1):64–76.

[Stolt, 1978] Stolt, R. H. (1978). Migration by Fourier transform. *Geophysics*, 1:23–48.

[Sun et al., 2002] Sun, J., Shum, H. Y., and Zheng, N. N. (2002). Stereo matching using belief propagation. In *Proceedings of the European Conference on Computer Vision*, volume 2, pages 510–524, Copenhagen, Denmark.

[Tomiyasu, 1978] Tomiyasu, K. (1978). Tutorial review of synthetic-aperture-radar *SAR* with applications to imaging of the ocean surface. *Proceedings of the IEEE*, 66(5):563–583.

[Touzi and Lopes, 1996] Touzi, R. and Lopes, A. (1996). Statistics of the Stokes parameters and of the complex coherence parameters in one-look and multi-look speckle field. *IEEE Trans. Geosci. Remote Sensing*, 34(2):519–532.

[Touzi et al., 1999] Touzi, R., Lopes, A., Bruniquel, J., and Vachon, P. W. (1999). Coherence estimation for SAR imagery. *IEEE Trans. Geoscience and Remote Sensing*, 37(1):135–149.

[Ulander et al., 2003] Ulander, L., Hellsten, H., and Stenström, G. (2003). Synthetic-aperture radar processing using fast factorised back-projection. *IEEE Trans. Aerosp. Electron. Syst.*, 39(3):760–776.

[Ulander and Frölind, 1998] Ulander, L. M. H. and Frölind, P. (1998). Ultra-wideband SAR interferometry. *IEEE Trans. Geoscience and Remote Sensing*, 36(5):1540–1550.

[Urick, 1975] Urick, R. J. (1975). *Principles of Underwater Sound*. M^{c}Graw-Hill.

[Vickery, 1998] Vickery, K. (1998). Acoustic positioning systems. a practical overview of current systems. In *Autonomous Underwater Vehicles, 1998. AUV'98. Proceedings Of The 1998 Workshop on*, pages 5–17.

[Wang et al., 2001] Wang, L., Bellettini, A., Hollett, R., Tesei, A., and Pinto, M. (2001). InSAS'00: Interferometric SAS and INS aided SAS imaging. In *Oceans 2001*, volume 1, pages 179–187. IEEE.

[Wilby, 1999] Wilby, A. (1999). The advantages, challenges and practical implementation of an interferometric swath bathymetry system. *IEEE J. Oceanic Engineering*, 1:23–29.

[Wu et al., 1982] Wu, C., Liu, K. Y., and Jin, M. (1982). Modeling and a correlation algorithm for spaceborne SAR signals. *IEEE Trans. Aerosp. Electron. Syst.*, 18(5):563–575.

[Xu et al., 1994] Xu, W., Chang, C., Kwoh, L. K., Lim, H., and Heng, W. C. A. (1994). Phase-unwrapping of SAR interferogram with multi-frequency or multi-baseline. In *Proceedings of the IEEE 1994 International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 730–732, Pasadena, CA.

[Xu and Stewart, 1999] Xu, W. and Stewart, W. K. (1999). Coherent source direction estimation fot[*sic*] three-row bathymetric sidescan sonars. In *Oceans 1999*, volume 1, pages 299–304.

[Zebker and Villasenor, 1992] Zebker, H. A. and Villasenor, J. (1992). Decorrelation in interferometric radar echos. *IEEE Trans. Geosci. Remote Sensing*, 30(5):950–959.

INDEX